

## (12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
26 June 2003 (26.06.2003)

PCT

(10) International Publication Number  
**WO 03/051912 A2**

- (51) International Patent Classification<sup>7</sup>: **C07K 14/18**
- (21) International Application Number: **PCT/EP02/14480**
- (22) International Filing Date:  
18 December 2002 (18.12.2002)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:  
10/020,510 18 December 2001 (18.12.2001) US  
60/418,358 16 October 2002 (16.10.2002) US
- (71) Applicant (for all designated States except US): **INNO-GENETICS N.V.** [BE/BE]; Technologiepark 6, B-9052 Ghent (BE).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): **MAERTENS, Geert** [BE/BE]; Zilverparrenstraat 64, B-8310 Brugge (BE). **DEPLA, Erik** [BE/BE]; Burgstraat 58, B-9070 Destelbergen (BE). **BOSMAN, Fons** [BE/BE]; Hulst 165, B-1745 Opwijk (BE).
- (74) Common Representative: **INNOGENETICS N.V.**; Intellectual Property Department, Industriepark Zwijnaarde 7, Box 4, B-9052 Ghent (BE).
- (81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- Published:**  
— without international search report and to be republished upon receipt of that report
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*



**WO 03/051912 A2**

(54) Title: **PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE**

(57) Abstract: The present invention relates to a method for purifying recombinant HCV single or specific oligomeric envelope proteins selected from the group consisting of E1 and/or E2 and/or E1/E2, characterized in that upon lysing the transformed host cells to isolate the recombinantly expressed protein a disulphide bond cleavage or reduction step is carried out with a disulphide bond cleavage agent. The present invention also relates to a composition isolated by such a method. The present invention also relates to the diagnostic and therapeutic application of these compositions. Furthermore, the invention relates to the use of HCV E1 protein and peptides for prognosing and monitoring the clinical effectiveness and/or clinical outcome of HCV treatment.

**PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC  
AND THERAPEUTIC USE**

**Field of the invention**

5           The present invention relates to the general fields of recombinant protein expression, purification of recombinant proteins, synthetic peptides, diagnosis of HCV infection, prophylactic treatment against HCV infection and to the prognosis/monitoring of the clinical efficiency of treatment of an individual with chronic hepatitis, or the prognosis/monitoring of natural disease.

10           More particularly, the present invention relates to purification methods for hepatitis C virus envelope proteins, the use in diagnosis, prophylaxis or therapy of HCV envelope proteins purified according to the methods described in the present invention, the use of single or specific oligomeric E1 and/or E2 and/or E1/E2 envelope proteins in assays for monitoring disease, and/or diagnosis of disease, and/or treatment of disease. The invention  
15 also relates to epitopes of the E1 and/or E2 envelope proteins and monoclonal antibodies thereto, as well their use in diagnosis, prophylaxis or treatment.

**Background of the invention**

20           The E2 protein purified from cell lysates according to the methods described in the present invention reacts with approximately 95% of patient sera. This reactivity is similar to the reactivity obtained with E2 secreted from CHO cells (Spaete et al., 1992). However, the intracellularly expressed form of E2 may more closely resemble the native viral envelope protein because it contains high mannose carbohydrate motifs, whereas the E2 protein secreted from CHO cells is further modified with galactose and sialic acid sugar moieties.  
25           When the aminoterminal half of E2 is expressed in the baculovirus system, only about 13 to 21% of sera from several patient groups can be detected (Inoue et al., 1992). After expression of E2 from E. coli, the reactivity of HCV sera was even lower and ranged from 14 (Yokosuka et al., 1992) to 17% (Mita et al., 1992). About 75% of HCV sera (and 95% of chronic patients) are anti-E1 positive using the purified, vaccinia-expressed recombinant E1  
30 protein of the present invention, in sharp contrast with the results of Kohara et al. (1992) and Hsu et al. (1993). Kohara et al. used a vaccinia-virus expressed E1 protein and detected anti-E1 antibodies in 7 to 23% of patients, while Hsu et al. only detected 14/50 (28%) sera using baculovirus-expressed E1.

These results show that not only a good expression system but also a good purification protocol are required to reach a high reactivity of the envelope proteins with human patient sera. This can be obtained using the proper expression system and/or purification protocols of the present invention which guarantee the conservation of the natural folding of the protein and the purification protocols of the present invention which guarantee the elimination of contaminating proteins and which preserve the conformation, and thus the reactivity of the HCV envelope proteins. The amounts of purified HCV envelope protein needed for diagnostic screening assays are in the range of grams per year. For vaccine purposes, even higher amounts of envelope protein would be needed. Therefore, the vaccinia virus system may be used for selecting the best expression constructs and for limited upscaling, and large-scale expression and purification of single or specific oligomeric envelope proteins containing high-mannose carbohydrates may be achieved when expressed from several yeast strains. In the case of hepatitis B for example, manufacturing of HBsAg from mammalian cells was much more costly compared with yeast-derived hepatitis B vaccines.

*Clinical importance of necro-inflammation and fibrosis in HCV infection.*

The natural history of liver disease after HCV infection does vary significantly from patient to patient. About 20% of the acutely infected persons are able to resolve infection spontaneously, while 80% of infected persons progresses to a chronic infection. Chronic infection results in an ongoing inflammation and/or necrosis (=necro-inflammation) in the liver which can be diagnosed by histological analysis of a liver biopsy or which can be diagnosed using a surrogate marker such as the presence of the liver enzyme ALT in serum. This chronic infection increases the risk for development of fibrosis which can lead to development of cirrhosis and ultimately liver carcinoma. Many data suggest that the ongoing necro-inflammation drives progression to fibrosis and cirrhosis. It is estimated that up to 20% of HCV chronic carriers may develop cirrhosis over a time period of about 20 years and that of those with cirrhosis between 1 to 4%/year is at risk to develop liver carcinoma. (Lauer and Walker 2001, Shiffman 1999). Both cirrhosis and liver carcinoma are end-stage liver diseases for which the treatment options are limited to liver transplantation. Consequently, the most important aim of therapy for HCV is to reduce the risk of development of end-stage liver disease by reducing liver necro-inflammation and/or reducing fibrosis progression. For the documentation and/or diagnosis of liver damage several scoring systems have been

developed for histological interpretation of a liver biopsy. These scoring systems may combine inflammation, necrosis and fibrosis in a single score such as the Histology Activity Index (HAI). Other scoring systems have separated the scores for necro-inflammation (=grading) from the one for fibrosis/cirrhosis (=staging). These systems  
5 include the system proposed by Ishak or the Metavir scoring system. A review of these scoring systems was published by Lefkowitz in 1997.

It has been shown in several studies that treatment with interferon, and more recently treatment with interferon combined with ribavirin and most recently treatment with pegylated interferon with or without ribavirin does change the natural history of HCV and halts further  
10 progression of liver fibrosis especially in those patients with a sustained viral response (Schvarcz et al. 1999, Shiffman 1999, Reichard et al. 1999, Poynard et al. 2002). The reduction of the risk for hepatocarcinogenesis in persons with sustained virological and even with sustained biochemical response has also been documented (Takimoto et al. 2002).

For persons without sustained virological response to interferon based therapy a  
15 maintenance interferon therapy may be helpful to prevent histological progression, but this only in a subset of patients (Alric et al. 2001).

Thus, many patients who do not respond to interferon based therapies or who are excluded from these therapies for several reasons (this may mount up to 70% of patients referred to clinic, Falck-Ytter et al. 2002), remain without a therapeutic option today to reduce liver  
20 necro-inflammation and/or reduce the progression of fibrosis in order to avoid end-stage liver disease.

#### Aims of the invention

It is an aim of the present invention to provide a new purification method for  
25 recombinantly expressed E1 and/or E2 and/or E1/E2 proteins such that said recombinant proteins are directly usable for diagnostic and vaccine purposes as single or specific oligomeric recombinant proteins free from contaminants instead of aggregates.

It is another aim of the present invention to provide compositions comprising purified (single or specific oligomeric) recombinant E1 and/or E2 and/or E1/E2 glycoproteins  
30 comprising conformational epitopes from the E1 and/or E2 domains of HCV.

It is yet another aim of the present invention to provide novel recombinant vector constructs for recombinantly expressing E1 and/or E2 and/or E1/E2 proteins, as well as host cells transformed with said vector constructs.

It is also an aim of the present invention to provide a method for producing and purifying recombinant HCV E1 and/or E2 and/or E1/E2 proteins.

It is also an aim of the present invention to provide diagnostic and immunogenic uses of the recombinant HCV E1 and/or E2 and/or E1/E2 proteins of the present invention, as well as to provide kits for diagnostic use, vaccines or therapeutics comprising any of the recombinant HCV E1 and/or E2 and/or E1/E2 proteins of the present invention.

It is further an aim of the present invention to provide for a new use of E1, E2, and/or E1/E2 proteins, or suitable parts thereof, for monitoring/prognosing the response to treatment of patients (e.g. with interferon) suffering from HCV infection.

It is also an aim of the present invention to provide for the use of the recombinant E1, E2, and/or E1/E2 proteins of the present invention in HCV screening and confirmatory antibody tests.

It is also an aim of the present invention to provide E1 and/or E2 peptides which can be used for diagnosis of HCV infection and for raising antibodies. Such peptides may also be used to isolate human monoclonal antibodies.

It is also an aim of the present invention to provide monoclonal antibodies, more particularly human monoclonal antibodies or mouse monoclonal antibodies which are humanized, which react specifically with E1 and/or E2 epitopes, either comprised in peptides or conformational epitopes comprised in recombinant proteins.

It is also an aim of the present invention to provide possible uses of anti-E1 or anti-E2 monoclonal antibodies for HCV antigen detection or for therapy of chronic HCV infection.

It is also an aim of the present invention to provide kits for monitoring/prognosing the response to treatment (e.g. with interferon) of patients suffering from HCV infection or monitoring/prognosing the outcome of the disease.

All the aims of the present invention are considered to have been met by the embodiments as set out below.

### Definitions

The following definitions serve to illustrate the different terms and expressions used in the present invention.

The term 'hepatitis C virus single envelope protein' refers to a polypeptide or an analogue thereof (e.g. mimotopes) comprising an amino acid sequence (and/or amino acid analogues) defining at least one HCV epitope of either the E1 or the E2 region. These single

envelope proteins in the broad sense of the word may be both monomeric or homo-oligomeric forms of recombinantly expressed envelope proteins. Typically, the sequences defining the epitope correspond to the amino acid sequence of either the E1 or the E2 region of HCV (either identically or via substitution of analogues of the native amino acid residue that do not destroy the epitope). In general, the epitope-defining sequence will be 3 or more amino acids in length, more typically, 5 or more amino acids in length, more typically 8 or more amino acids in length, and even more typically 10 or more amino acids in length. With respect to conformational epitopes, the length of the epitope-defining sequence can be subject to wide variations, since it is believed that these epitopes are formed by the three-dimensional shape of the antigen (e.g. folding). Thus, the amino acids defining the epitope can be relatively few in number, but widely dispersed along the length of the molecule being brought into the correct epitope conformation via folding. The portions of the antigen between the residues defining the epitope may not be critical to the conformational structure of the epitope. For example, deletion or substitution of these intervening sequences may not affect the conformational epitope provided sequences critical to epitope conformation are maintained (e.g. cysteines involved in disulfide bonding, glycosylation sites, etc.). A conformational epitope may also be formed by 2 or more essential regions of subunits of a homooligomer or heterooligomer.

The HCV antigens of the present invention comprise conformational epitopes from the E1 and/or E2 (envelope) domains of HCV. The E1 domain, which is believed to correspond to the viral envelope protein, is currently estimated to span amino acids 192-383 of the HCV polyprotein (Hijikata et al., 1991). Upon expression in a mammalian system (glycosylated), it is believed to have an approximate molecular weight of 35 kDa as determined via SDS-PAGE. The E2 protein, previously called NS1, is believed to span amino acids 384-809 or 384-746 (Grakoui et al., 1993) of the HCV polyprotein and to also be an envelope protein. Upon expression in a vaccinia system (glycosylated), it is believed to have an apparent gel molecular weight of about 72 kDa. It is understood that these protein endpoints are approximations (e.g. the carboxy terminal end of E2 could lie somewhere in the 730-820 amino acid region, e.g. ending at amino acid 730, 735, 740, 742, 744, 745, preferably 746, 747, 748, 750, 760, 770, 780, 790, 800, 809, 810, 820). The E2 protein may also be expressed together with the E1, P7 (aa 747-809), NS2 (aa 810-1026), NS4A (aa 1658-1711) or NS4B (aa 1712-1972). Expression together with these other HCV proteins may be important for obtaining the correct protein folding.

It is also understood that the isolates used in the examples section of the present invention were not intended to limit the scope of the invention and that any HCV isolate from type 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or any other new genotype of HCV is a suitable source of E1 and/or E2 sequence for the practice of the present invention.

5       The E1 and E2 antigens used in the present invention may be full-length viral proteins, substantially full-length versions thereof, or functional fragments thereof (e.g. fragments which are not missing sequence essential to the formation or retention of an epitope). Furthermore, the HCV antigens of the present invention can also include other sequences that do not block or prevent the formation of the conformational epitope of  
10       interest. The presence or absence of a conformational epitope can be readily determined though screening the antigen of interest with an antibody (polyclonal serum or monoclonal to the conformational epitope) and comparing its reactivity to that of a denatured version of the antigen which retains only linear epitopes (if any). In such screening using polyclonal antibodies, it may be advantageous to adsorb the polyclonal serum first with the denatured  
15       antigen and see if it retains antibodies to the antigen of interest.

      The HCV antigens of the present invention can be made by any recombinant method that provides the epitope of interest. For example, recombinant intracellular expression in mammalian or insect cells is a preferred method to provide glycosylated E1 and/or E2 antigens in 'native' conformation as is the case for the natural HCV antigens. Yeast cells and  
20       mutant yeast strains (e.g. mnn 9 mutant (Kniskern et al., 1994) or glycosylation mutants derived by means of vanadate resistance selection (Ballou et al., 1991)) may be ideally suited for production of secreted high-mannose-type sugars; whereas proteins secreted from mammalian cells may contain modifications including galactose or sialic acids which may be undesirable for certain diagnostic or vaccine applications. However, it may also be possible  
25       and sufficient for certain applications, as it is known for proteins, to express the antigen in other recombinant hosts (such as *E. coli*) and renature the protein after recovery.

      The term 'fusion polypeptide' intends a polypeptide in which the HCV antigen(s) are part of a single continuous chain of amino acids, which chain does not occur in nature. The HCV antigens may be connected directly to each other by peptide bonds or be separated by  
30       intervening amino acid sequences. The fusion polypeptides may also contain amino acid sequences exogenous to HCV.

      The term 'solid phase' intends a solid body to which the individual HCV antigens or the fusion polypeptide comprised of HCV antigens are bound covalently or by noncovalent

means such as hydrophobic adsorption.

The term 'biological sample' intends a fluid or tissue of a mammalian individual (e.g. an anthropoid, a human) that commonly contains antibodies produced by the individual, more particularly antibodies against HCV. The fluid or tissue may also contain HCV antigen.

5 Such components are known in the art and include, without limitation, blood, plasma, serum, urine, spinal fluid, lymph fluid, secretions of the respiratory, intestinal or genitourinary tracts, tears, saliva, milk, white blood cells and myelomas. Body components include biological liquids. The term 'biological liquid' refers to a fluid obtained from an organism. Some biological fluids are used as a source of other products, such as clotting factors (e.g. Factor

10 VIII;C), serum albumin, growth hormone and the like. In such cases, it is important that the source of biological fluid be free of contamination by virus such as HCV.

The term 'immunologically reactive' means that the antigen in question will react specifically with anti-HCV antibodies present in a body component from an HCV infected individual.

15 The term 'immune complex' intends the combination formed when an antibody binds to an epitope on an antigen.

'E1' as used herein refers to a protein or polypeptide expressed within the first 400 amino acids of an HCV polyprotein, sometimes referred to as the E, ENV or S protein. In its natural form it is a 35 kDa glycoprotein which is found in strong association with

20 membranes. In most natural HCV strains, the E1 protein is encoded in the viral polyprotein following the C (core) protein. The E1 protein extends from approximately amino acid (aa) 192 to about aa 383 of the full-length polyprotein.

The term 'E1' as used herein also includes analogs and truncated forms that are immunologically cross-reactive with natural E1, and includes E1 proteins of genotypes 1, 2,

25 3, 4, 5, 6, 7, 8, 9, 10, or any other newly identified HCV type or subtype.

'E2' as used herein refers to a protein or polypeptide expressed within the first 900 amino acids of an HCV polyprotein, sometimes referred to as the NS1 protein. In its natural form it is a 72 kDa glycoprotein that is found in strong association with membranes. In most natural HCV strains, the E2 protein is encoded in the viral polyprotein following the E1

30 protein. The E2 protein extends from approximately amino acid position 384 to amino acid position 746, another form of E2 extends to amino acid position 809. The term 'E2' as used herein also includes analogs and truncated forms that are immunologically cross-reactive with natural E2. For example, insertions of multiple codons between codon 383 and 384, as



well as deletions of amino acids 384-387 have been reported by Kato et al. (1992).

'E1/E2' as used herein refers to an oligomeric form of envelope proteins containing at least one E1 component and at least one E2 component.

The term 'specific oligomeric' E1 and/or E2 and/or E1/E2 envelope proteins refers to  
5 all possible oligomeric forms of recombinantly expressed E1 and/or E2 envelope proteins which are not aggregates. E1 and/or E2 specific oligomeric envelope proteins are also referred to as homo-oligomeric E1 or E2 envelope proteins (see below).

The term 'single or specific oligomeric' E1 and/or E2 and/or E1/E2 envelope proteins refers to single monomeric E1 or E2 proteins (single in the strict sense of the word) as well as  
10 specific oligomeric E1 and/or E2 and/or E1/E2 recombinantly expressed proteins. These single or specific oligomeric envelope proteins according to the present invention can be further defined by the following formula  $(E1)_x(E2)_y$ , wherein x can be a number between 0 and 100, and y can be a number between 0 and 100, provided that x and y are not both 0. With x=1 and y=0 said envelope proteins include monomeric E1.

The term 'homo-oligomer' as used herein refers to a complex of E1 and/or E2  
15 containing more than one E1 or E2 monomer, e.g. E1/E1 dimers, E1/E1/E1 trimers or E1/E1/E1/E1 tetramers and E2/E2 dimers, E2/E2/E2 trimers or E2/E2/E2/E2 tetramers, E1 pentamers and hexamers, E2 pentamers and hexamers or any higher-order homo-oligomers of E1 or E2 are all 'homo-oligomers' within the scope of this definition. The oligomers may  
20 contain one, two, or several different monomers of E1 or E2 obtained from different types or subtypes of hepatitis C virus including for example those described in an international application published under WO 94/25601 and European application No. 94870166.9 both by the present applicants. Such mixed oligomers are still homo-oligomers within the scope of this invention, and may allow more universal diagnosis, prophylaxis or treatment of HCV.

The term 'purified' as applied to proteins herein refers to a composition wherein the  
25 desired protein comprises at least 35% of the total protein component in the composition. The desired protein preferably comprises at least 40%, more preferably at least about 50%, more preferably at least about 60%, still more preferably at least about 70%, even more preferably at least about 80%, even more preferably at least about 90%, and most preferably  
30 at least about 95% of the total protein component. The composition may contain other compounds such as carbohydrates, salts, lipids, solvents, and the like, without affecting the determination of the percentage purity as used herein. An 'isolated' HCV protein intends an HCV protein composition that is at least 35% pure.

The term 'essentially purified proteins' refers to proteins purified such that they can be used for in vitro diagnostic methods and as a therapeutic compound. These proteins are substantially free from cellular proteins, vector-derived proteins or other HCV viral components. Usually these proteins are purified to homogeneity (at least 80% pure, preferably, 90%, more preferably 95%, more preferably 97%, more preferably 98%, more preferably 99%, even more preferably 99.5%, and most preferably the contaminating proteins should be undetectable by conventional methods like SDS-PAGE and silver staining.

The term 'recombinantly expressed' used within the context of the present invention refers to the fact that the proteins of the present invention are produced by recombinant expression methods be it in prokaryotes, or lower or higher eukaryotes as discussed in detail below.

The term 'lower eukaryote' refers to host cells such as yeast, fungi and the like. Lower eukaryotes are generally (but not necessarily) unicellular. Preferred lower eukaryotes are yeasts, particularly species within *Saccharomyces*, *Schizosaccharomyces*, *Kluyveromyces*, *Pichia* (e.g. *Pichia pastoris*), *Hansenula* (e.g. *Hansenula polymorpha*), *Yarrowia*, *Schwaniomyces*, *Schizosaccharomyces*, *Zygosaccharomyces* and the like. *Saccharomyces cerevisiae*, *S. carlsbergensis* and *K. lactis* are the most commonly used yeast hosts, and are convenient fungal hosts.

The term 'prokaryotes' refers to hosts such as *E. coli*, *Lactobacillus*, *Lactococcus*, *Salmonella*, *Streptococcus*, *Bacillus subtilis* or *Streptomyces*. Also these hosts are contemplated within the present invention.

The term 'higher eukaryote' refers to host cells derived from higher animals, such as mammals, reptiles, insects, and the like. Presently preferred higher eukaryote host cells are derived from Chinese hamster (e.g. CHO), monkey (e.g. COS and Vero cells), baby hamster kidney (BHK), pig kidney (PK15), rabbit kidney 13 cells (RK13), the human osteosarcoma cell line 143 B, the human cell line HeLa and human hepatoma cell lines like Hep G2, and insect cell lines (e.g. *Spodoptera frugiperda*). The host cells may be provided in suspension or flask cultures, tissue cultures, organ cultures and the like. Alternatively the host cells may also be transgenic animals.

The term 'polypeptide' refers to a polymer of amino acids and does not refer to a specific length of the product; thus, peptides, oligopeptides, and proteins are included within the definition of polypeptide. This term also does not refer to or exclude post-expression modifications of the polypeptide, for example, glycosylations, acetylations, phosphorylations

and the like. Included within the definition are, for example, polypeptides containing one or more analogues of an amino acid (including, for example, unnatural amino acids, PNA, etc.), polypeptides with substituted linkages, as well as other modifications known in the art, both naturally occurring and non-naturally occurring.

5           The term 'recombinant polynucleotide or nucleic acid' intends a polynucleotide or nucleic acid of genomic, cDNA, semisynthetic, or synthetic origin which, by virtue of its origin or manipulation : (1) is not associated with all or a portion of a polynucleotide with which it is associated in nature, (2) is linked to a polynucleotide other than that to which it is linked in nature, or (3) does not occur in nature.

10           The term 'recombinant host cells', 'host cells', 'cells', 'cell lines', 'cell cultures', and other such terms denoting microorganisms or higher eukaryotic cell lines cultured as unicellular entities refer to cells which can be or have been, used as recipients for a recombinant vector or other transfer polynucleotide, and include the progeny of the original cell which has been transfected. It is understood that the progeny of a single parental cell may  
15           not necessarily be completely identical in morphology or in genomic or total DNA complement as the original parent, due to natural, accidental, or deliberate mutation.

          The term 'replicon' is any genetic element, e.g., a plasmid, a chromosome, a virus, a cosmid, etc., that behaves as an autonomous unit of polynucleotide replication within a cell; i.e., capable of replication under its own control.

20           The term 'vector' is a replicon further comprising sequences providing replication and/or expression of a desired open reading frame.

          The term 'control sequence' refers to polynucleotide sequences which are necessary to effect the expression of coding sequences to which they are ligated. The nature of such control sequences differs depending upon the host organism; in prokaryotes, such control  
25           sequences generally include promoter, ribosomal binding site, and terminators; in eukaryotes, generally, such control sequences include promoters, terminators and, in some instances, enhancers. The term 'control sequences' is intended to include, at a minimum, all components whose presence is necessary for expression, and may also include additional components whose presence is advantageous, for example, leader sequences which govern secretion.

30           The term 'promoter' is a nucleotide sequence which is comprised of consensus sequences which allow the binding of RNA polymerase to the DNA template in a manner such that mRNA production initiates at the normal transcription initiation site for the adjacent structural gene.

The expression 'operably linked' refers to a juxtaposition wherein the components so described are in a relationship permitting them to function in their intended manner. A control sequence 'operably linked' to a coding sequence is ligated in such a way that expression of the coding sequence is achieved under conditions compatible with the control sequences.

An 'open reading frame' (ORF) is a region of a polynucleotide sequence which encodes a polypeptide and does not contain stop codons; this region may represent a portion of a coding sequence or a total coding sequence.

A 'coding sequence' is a polynucleotide sequence which is transcribed into mRNA and/or translated into a polypeptide when placed under the control of appropriate regulatory sequences. The boundaries of the coding sequence are determined by a translation start codon at the 5'-terminus and a translation stop codon at the 3'-terminus. A coding sequence can include but is not limited to mRNA, DNA (including cDNA), and recombinant polynucleotide sequences.

As used herein, 'epitope' or 'antigenic determinant' means an amino acid sequence that is immunoreactive. Generally an epitope consists of at least 3 to 4 amino acids, and more usually, consists of at least 5 or 6 amino acids, sometimes the epitope consists of about 7 to 8, or even about 10 amino acids. As used herein, an epitope of a designated polypeptide denotes epitopes with the same amino acid sequence as the epitope in the designated polypeptide, and immunologic equivalents thereof. Such equivalents also include strain, subtype (=genotype), or type(group)-specific variants, e.g. of the currently known sequences or strains belonging to genotypes 1a, 1b, 1c, 1d, 1e, 1f, 2a, 2b, 2c, 2d, 2e, 2f, 2g, 2h, 2i, 3a, 3b, 3c, 3d, 3e, 3f, 3g, 4a, 4b, 4c, 4d, 4e, 4f, 4g, 4h, 4i, 4j, 4k, 4l, 5a, 5b, 6a, 6b, 6c, 7a, 7b, 7c, 8a, 8b, 9a, 9b, 10a, or any other newly defined HCV (sub)type. It is to be understood that the amino acids constituting the epitope need not be part of a linear sequence, but may be interspersed by any number of amino acids, thus forming a conformational epitope.

The term 'immunogenic' refers to the ability of a substance to cause a humoral and/or cellular response, whether alone or when linked to a carrier, in the presence or absence of an adjuvant. 'Neutralization' refers to an immune response that blocks the infectivity, either partially or fully, of an infectious agent. A 'vaccine' is an immunogenic composition capable of eliciting protection against HCV, whether partial or complete. A vaccine may also be useful for treatment of an individual, in which case it is called a therapeutic vaccine.

The term 'therapeutic' refers to a composition capable of treating HCV infection.

The term 'effective amount' refers to an amount of epitope-bearing polypeptide sufficient to induce an immunogenic response in the individual to which it is administered, or to otherwise detectably immunoreact in its intended system (e.g., immunoassay). Preferably, the effective amount is sufficient to effect treatment, as defined above. The exact amount  
5 necessary will vary according to the application. For vaccine applications or for the generation of polyclonal antiserum / antibodies, for example, the effective amount may vary depending on the species, age, and general condition of the individual, the severity of the condition being treated, the particular polypeptide selected and its mode of administration, etc. It is also believed that effective amounts will be found within a relatively large, non-  
10 critical range. An appropriate effective amount can be readily determined using only routine experimentation. Preferred ranges of E1 and/or E2 and/or E1/E2 single or specific oligomeric envelope proteins for prophylaxis of HCV disease are 0.01 to 100 µg/dose, preferably 0.1 to 50 µg/dose. Several doses may be needed per individual in order to achieve a sufficient immune response and subsequent protection against HCV disease.

#### Detailed description of the invention

More particularly, the present invention contemplates a method for isolating or purifying recombinant HCV single or specific oligomeric envelope protein selected from the group consisting of E1 and/or E2 and/or E1/E2, characterized in that upon lysing the  
20 transformed host cells to isolate the recombinantly expressed protein a disulphide bond cleavage or reduction step is carried out with a disulphide bond cleaving agent.

The essence of these 'single or specific oligomeric' envelope proteins of the invention is that they are free from contaminating proteins and that they are not disulphide bond linked with contaminants.

25 The proteins according to the present invention are recombinantly expressed in lower or higher eukaryotic cells or in prokaryotes. The recombinant proteins of the present invention are preferably glycosylated and may contain high-mannose-type, hybrid, or complex glycosylations. Preferentially said proteins are expressed from mammalian cell lines as discussed in detail in the Examples section, or in yeast such as in mutant yeast strains also  
30 as detailed in the Examples section.

The proteins according to the present invention may be secreted or expressed within components of the cell, such as the ER or the Golgi Apparatus. Preferably, however, the proteins of the present invention bear high-mannose-type glycosylations and are retained in

the ER or Golgi Apparatus of mammalian cells or are retained in or secreted from yeast cells, preferably secreted from yeast mutant strains such as the *mn9* mutant (Kniskern et al., 1994), or from mutants that have been selected by means of vanadate resistance (Ballou et al., 1991).

5        Upon expression of HCV envelope proteins, the present inventors could show that some of the free thiol groups of cysteines not involved in intra- or inter-molecular disulphide bridges, react with cysteines of host or expression-system-derived (e.g. vaccinia) proteins or of other HCV envelope proteins (single or oligomeric), and form aspecific intermolecular bridges. This results in the formation of 'aggregates' of HCV envelope proteins together with  
10        contaminating proteins. It was also shown in WO 92/08734 that 'aggregates' were obtained after purification, but it was not described which protein interactions were involved. In patent application WO 92/08734, recombinant E1/E2 protein expressed with the vaccinia virus system were partially purified as aggregates and only found to be 70% pure, rendering the purified aggregates not useful for diagnostic, prophylactic or therapeutic purposes.

15        Therefore, a major aim of the present invention resides in the separation of single or specific-oligomeric HCV envelope proteins from contaminating proteins, and to use the purified proteins (> 95% pure) for diagnostic, prophylactic and therapeutic purposes. To those purposes, the present inventors have been able to provide evidence that aggregated protein complexes ('aggregates') are formed on the basis of disulphide bridges and non-  
20        covalent protein-protein interactions. The present invention thus provides a means for selectively cleaving the disulphide bonds under specific conditions and for separating the cleaved proteins from contaminating proteins which greatly interfere with diagnostic, prophylactic and therapeutic applications. The free thiol groups may be blocked (reversibly or irreversibly) in order to prevent the reformation of disulphide bridges, or may be left to  
25        oxidize and oligomerize with other envelope proteins (see definition homo-oligomer). It is to be understood that such protein oligomers are essentially different from the 'aggregates' described in WO 92/08734 and WO 94/01778, since the level of contaminating proteins is undetectable.

Said disulphide bond cleavage may also be achieved by:

- 30        (1) performic acid oxidation by means of cysteic acid in which case the cysteine residues are modified into cysteic acid (Moore et al., 1963).  
      (2) Sulfitolysis ( $R-S-S-R \rightarrow 2 R-SO_3^-$ ) for example by means of sulphite ( $SO_3^{2-}$ ) together with a proper oxidant such as  $Cu^{2+}$  in which case the cysteine is modified into S-sulpho-cysteine

(Bailey and Cole, 1959).

(3) Reduction by means of mercaptans, such as dithiotreitol (DDT),  $\beta$ -mercapto-ethanol, cysteine, glutathione Red,  $\epsilon$ -mercapto-ethylamine, or thioglycolic acid, of which DTT and  $\beta$ -mercapto-ethanol are commonly used (Cleland, 1964), is the preferred method of this invention because the method can be performed in a water environment and because the cysteine remains unmodified.

(4) Reduction by means of a phosphine (e.g.  $\text{Bu}_3\text{P}$ ) (Ruegg and Rudinger, 1977).

All these compounds are thus to be regarded as agents or means for cleaving disulphide bonds according to the present invention.

Said disulphide bond cleavage (or reducing) step of the present invention is preferably a partial disulphide bond cleavage (reducing) step (carried out under partial cleavage or reducing conditions).

A preferred disulphide bond cleavage or reducing agent according to the present invention is dithiothreitol (DTT). Partial reduction is obtained by using a low concentration of said reducing agent, i.e. for DTT for example in the concentration range of about 0.1 to about 50 mM, preferably about 0.1 to about 20 mM, preferably about 0.5 to about 10 mM, preferably more than 1 mM, more than 2 mM or more than 5 mM, more preferably about 1.5 mM, about 2.0 mM, about 2.5 mM, about 5 mM or about 7.5 mM.

Said disulphide bond cleavage step may also be carried out in the presence of a suitable detergent (as an example of a means for cleaving disulphide bonds or in combination with a cleaving agent) able to dissociate the expressed proteins, such as DecylPEG, EMPIGEN-BB, NP-40, sodium cholate, Triton X-100.

Said reduction or cleavage step (preferably a partial reduction or cleavage step) is carried out preferably in the presence of (with) a detergent. A preferred detergent according to the present invention is Empigen-BB. The amount of detergent used is preferably in the range of 1 to 10 %, preferably more than 3%, more preferably about 3.5% of a detergent such as Empigen-BB.

A particularly preferred method for obtaining disulphide bond cleavage employs a combination of a classical disulphide bond cleavage agent as detailed above and a detergent (also as detailed above). As contemplated in the Examples section, the particular combination of a low concentration of DTT (1.5 to 7.5 mM) and about 3.5 % of Empigen-BB is proven to be a particularly preferred combination of reducing agent and detergent for the purification of recombinantly expressed E1 and E2 proteins. Upon gelfiltration chromatography, said partial

reduction is shown to result in the production of possibly dimeric E1 protein and separation of this E1 protein from contaminating proteins that cause false reactivity upon use in immunoassays.

It is, however, to be understood that also any other combination of any reducing agent  
 5 known in the art with any detergent or other means known in the art to make the cysteines better accessible is also within the scope of the present invention, insofar as said combination reaches the same goal of disulphide bridge cleavage as the preferred combination exemplified in the present invention.

Apart from reducing the disulphide bonds, a disulphide bond cleaving means  
 10 according to the present invention may also include any disulphide bridge exchanging agents (competitive agent being either organic or proteinaeous, see for instance Creighton, 1988) known in the art which allows the following type of reaction to occur:



\* R1, R2: compounds of protein aggregates

15 \* R3 SH: competitive agent (organic, proteinaeous)

The term 'disulphide bridge exchanging agent' is to be interpreted as including disulphide bond reforming as well as disulphide bond blocking agents.

The present invention also relates to methods for purifying or isolating HCV single or specific oligomeric envelope proteins as set out above further including the use of any SH  
 20 group blocking or binding reagent known in the art such as chosen from the following list:

- Glutathion
- 5,5'-dithiobis-(2-nitrobenzoic acid) or bis-(3-carboxy-4-nitrophenyl)-disulphide (DTNB or Ellman's reagent) (Elmann, 1959)
- N-ethylmaleimide (NEM; Benesch et al., 1956)
- 25 - N-(4-dimethylamino-3,5-dinitrophenyl) maleimide or Tuppy's maleimide which provides a color to the protein
- P-chloromercuribenzoate (Grassetti et al., 1969)
- 4-vinylpyridine (Friedman and Krull, 1969) can be liberated after reaction by acid hydrolysis
- 30 - acrylonitrile, can be liberated after reaction by acid hydrolysis (Weil and Seibles, 1961)
- NEM-biotin (e.g. obtained from Sigma B1267)
- 2,2'-dithiopyridine (Grassetti and Murray, 1967)



- 4,4'-dithiopyridine (Grassetti and Murray, 1967)
- 6,6'-dithiodinicotinic acid (DTDNA; Brown and Cunningham, 1970)
- 2,2'-dithiobis-(5'-nitropyridine) (DTNP; US patent 3597160) or other dithiobis (heterocyclic derivative) compounds (Grassetti and Murray, 1969)

5 A survey of the publications cited shows that often different reagents for sulphhydryl groups will react with varying numbers of thiol groups of the same protein or enzyme molecule. One may conclude that this variation in reactivity of the thiol groups is due to the steric environment of these groups, such as the shape of the molecule and the surrounding groups of atoms and their charges, as well as to the size, shape and charge of the reagent  
10 molecule or ion. Frequently the presence of adequate concentrations of denaturants such as sodium dodecylsulfate, urea or guanidine hydrochloride will cause sufficient unfolding of the protein molecule to permit equal access to all of the reagents for thiol groups. By varying the concentration of denaturant, the degree of unfolding can be controlled and in this way thiol groups with different degrees of reactivity may be revealed. Although up to date most of the  
15 work reported has been done with p-chloromercuribenzoate, N-ethylmaleimide and DTNB, it is likely that the other more recently developed reagents may prove equally useful. Because of their varying structures, it seems likely, in fact, that they may respond differently to changes in the steric environment of the thiol groups.

Alternatively, conditions such as low pH (preferably lower than pH 6) for preventing  
20 free SH groups from oxidizing and thus preventing the formation of large intermolecular aggregates upon recombinant expression and purification of E1 and E2 (envelope) proteins are also within the scope of the present invention.

A preferred SH group blocking reagent according to the present invention is N-ethylmaleimide (NEM). Said SH group blocking reagent may be administered during lysis  
25 of the recombinant host cells and after the above-mentioned partial reduction process or after any other process for cleaving disulphide bridges. Said SH group blocking reagent may also be modified with any group capable of providing a detectable label and/or any group aiding in the immobilization of said recombinant protein to a solid substrate, e.g. biotinylated NEM.

Methods for cleaving cysteine bridges and blocking free cysteines have also been  
30 described in Darbre (1987), Means and Feeney (1971), and by Wong (1993).

A method to purify single or specific oligomeric recombinant E1 and/or E2 and/or E1/E2 proteins according to the present invention as defined above is further characterized as comprising the following steps:

- lysing recombinant E1 and/or E2 and/or E1/E2 expressing host cells, preferably in the presence of an SH group blocking agent, such as N-ethylmaleimide (NEM), and possibly a suitable detergent, preferably Empigen-BB,
- recovering said HCV envelope protein by affinity purification for instance by means  
5 lectin-chromatography, such as lentil-lectin chromatography, or immunoaffinity chromatography using anti-E1 and/or anti-E2 specific monoclonal antibodies, followed by,
- reduction or cleavage of disulphide bonds with a disulphide bond cleaving agent, such as DTT, preferably also in the presence of an SH group blocking agent, such as  
10 NEM or Biotin-NEM, and,
- recovering the reduced HCV E1 and/or E2 and/or E1/E2 envelope proteins for instance by gelfiltration (size exclusion chromatography or molecular sieving) and possibly also by an additional Ni<sup>2+</sup>-IMAC chromatography and desalting step.

It is to be understood that the above-mentioned recovery steps may also be carried out  
15 using any other suitable technique known by the person skilled in the art.

Preferred lectin-chromatography systems include *Galanthus nivalis* agglutinin (GNA) - chromatography, or *Lens culinaris* agglutinin (LCA) (lentil) lectin chromatography as illustrated in the Examples section. Other useful lectins include those recognizing high-mannose type sugars, such as *Narcissus pseudonarcissus* agglutinin (NPA), *Pisum sativum*  
20 agglutinin (PSA), or *Allium ursinum* agglutinin (AUA).

Preferably said method is usable to purify single or specific oligomeric HCV envelope protein produced intracellularly as detailed above.

For secreted E1 or E2 or E1/E2 oligomers, lectins binding complex sugars such as *Ricinus communis* agglutinin I (RCA I), are preferred lectins.

25 The present invention more particularly contemplates essentially purified recombinant HCV single or specific oligomeric envelope proteins, selected from the group consisting of E1 and/or E2 and/or E1/E2, characterized as being isolated or purified by a method as defined above.

The present invention more particularly relates to the purification or isolation of  
30 recombinant envelope proteins which are expressed from recombinant mammalian cells such as vaccinia.

The present invention also relates to the purification or isolation of recombinant envelope proteins which are expressed from recombinant yeast cells.

The present invention equally relates to the purification or isolation of recombinant envelope proteins which are expressed from recombinant bacterial (prokaryotic) cells.

The present invention also contemplates a recombinant vector comprising a vector sequence, an appropriate prokaryotic, eukaryotic or viral or synthetic promoter sequence  
5 followed by a nucleotide sequence allowing the expression of the single or specific oligomeric E1 and/or E2 and/or E1/E2 of the invention.

Particularly, the present invention contemplates a recombinant vector comprising a vector sequence, an appropriate prokaryotic, eukaryotic or viral or synthetic promoter sequence followed by a nucleotide sequence allowing the expression of the single E1 or E1  
10 of the invention.

Particularly, the present invention contemplates a recombinant vector comprising a vector sequence, an appropriate prokaryotic, eukaryotic or viral or synthetic promoter sequence followed by a nucleotide sequence allowing the expression of the single E1 or E2 of the invention.

15 The segment of the HCV cDNA encoding the desired E1 and/or E2 sequence inserted into the vector sequence may be attached to a signal sequence. Said signal sequence may be that from a non-HCV source, e.g. the IgG or tissue plasminogen activator (tpa) leader sequence for expression in mammalian cells, or the  $\alpha$ -mating factor sequence for expression into yeast cells, but particularly preferred constructs according to the present invention  
20 contain signal sequences appearing in the HCV genome before the respective start points of the E1 and E2 proteins. The segment of the HCV cDNA encoding the desired E1 and/or E2 sequence inserted into the vector may also include deletions e.g. of the hydrophobic domain(s) as illustrated in the examples section, or of the E2 hypervariable region I.

More particularly, the recombinant vectors according to the present invention  
25 encompass a nucleic acid having an HCV cDNA segment encoding the polyprotein starting in the region between amino acid positions 1 and 192 and ending in the region between positions 250 and 400 of the HCV polyprotein, more preferably ending in the region between positions 250 and 341, even more preferably ending in the region between positions 290 and 341 for expression of the HCV single E1 protein. Most preferably, the present recombinant  
30 vector encompasses a recombinant nucleic acid having a HCV cDNA segment encoding part of the HCV polyprotein starting in the region between positions 117 and 192, and ending at any position in the region between positions 263 and 326, for expression of HCV single E1 protein. Also within the scope of the present invention are forms that have the first

hydrophobic domain deleted (positions 264 to 293 plus or minus 8 amino acids), or forms to which a 5'-terminal ATG codon and a 3'-terminal stop codon has been added, or forms which have a factor Xa cleavage site and/or 3 to 10, preferably 6 Histidine codons have been added.

More particularly, the recombinant vectors according to the present invention  
5 encompass a nucleic acid having an HCV cDNA segment encoding the polyprotein starting in the region between amino acid positions 290 and 406 and ending in the region between positions 600 and 820 of the HCV polyprotein, more preferably starting in the region between positions 322 and 406, even more preferably starting in the region between positions 347 and 406, even still more preferably starting in the region between positions 364 and 406  
10 for expression of the HCV single E2 protein. Most preferably, the present recombinant vector encompasses a recombinant nucleic acid having a HCV cDNA segment encoding the polyprotein starting in the region between positions 290 and 406, and ending at any position of positions 623, 650, 661, 673, 710, 715, 720, 746 or 809, for expression of HCV single E2 protein. Also within the scope of the present invention are forms to which a 5'-terminal ATG  
15 codon and a 3'-terminal stop codon has been added, or forms which have a factor Xa cleavage site and/or 3 to 10, preferably 6 Histidine codons have been added.

A variety of vectors may be used to obtain recombinant expression of HCV single or specific oligomeric envelope proteins of the present invention. Lower eukaryotes such as yeasts and glycosylation mutant strains are typically transformed with plasmids, or are  
20 transformed with a recombinant virus. The vectors may replicate within the host independently, or may integrate into the host cell genome.

Higher eukaryotes may be transformed with vectors, or may be infected with a recombinant virus, for example a recombinant vaccinia virus. Techniques and vectors for the insertion of foreign DNA into vaccinia virus are well known in the art, and utilize, for  
25 example homologous recombination. A wide variety of viral promoter sequences, possibly terminator sequences and poly(A)-addition sequences, possibly enhancer sequences and possibly amplification sequences, all required for the mammalian expression, are available in the art. Vaccinia is particularly preferred since vaccinia halts the expression of host cell proteins. Vaccinia is also very much preferred since it allows the expression of E1 and E2  
30 proteins of HCV in cells or individuals which are immunized with the live recombinant vaccinia virus. For vaccination of humans the avipox and Ankara Modified Virus (AMV) are particularly useful vectors.

Also known are insect expression transfer vectors derived from baculovirus

*Autographa californica* nuclear polyhedrosis virus (AcNPV), which is a helper-independent viral expression vector. Expression vectors derived from this system usually use the strong viral polyhedrin gene promoter to drive the expression of heterologous genes. Different vectors as well as methods for the introduction of heterologous DNA into the desired site of baculovirus are available to the man skilled in the art for baculovirus expression. Also different signals for posttranslational modification recognized by insect cells are known in the art.

Also included within the scope of the present invention is a method for producing purified recombinant single or specific oligomeric HCV E1 or E2 or E1/E2 proteins, wherein the cysteine residues involved in aggregates formation are replaced at the level of the nucleic acid sequence by other residues such that aggregate formation is prevented. The recombinant proteins expressed by recombinant vectors carrying such a mutated E1 and/or E2 protein encoding nucleic acid are also within the scope of the present invention.

The present invention also relates to recombinant E1 and/or E2 and/or E1/E2 proteins characterized in that at least one of their glycosylation sites has been removed and are consequently termed glycosylation mutants. As explained in the Examples section, different glycosylation mutants may be desired to diagnose (screening, confirmation, prognosis, etc.) and prevent HCV disease according to the patient in question. An E2 protein glycosylation mutant lacking the GLY4 has for instance been found to improve the reactivity of certain sera in diagnosis. These glycosylation mutants are preferably purified according to the method disclosed in the present invention. Also contemplated within the present invention are recombinant vectors carrying the nucleic acid insert encoding such a E1 and/or E2 and/or E1/E2 glycosylation mutant as well as host cells transformed with such a recombinant vector.

The present invention also relates to recombinant vectors including a polynucleotide which also forms part of the present invention. The present invention relates more particularly to the recombinant nucleic acids as represented in SEQ ID NO 3, 5, 7, 9, 11, 13, 21, 23, 25, 27, 29, 31, 35, 37, 39, 41, 43, 45, 47 and 49, or parts thereof.

The present invention also contemplates host cells transformed with a recombinant vector as defined above, wherein said vector comprises a nucleotide sequence encoding HCV E1 and/or E2 and/or E1/E2 protein as defined above in addition to a regulatory sequence operably linked to said HCV E1 and/or E2 and/or E1/E2 sequence and capable of regulating the expression of said HCV E1 and/or E2 and/or E1/E2 protein.

Eukaryotic hosts include lower and higher eukaryotic hosts as described in the definitions section. Lower eukaryotic hosts include yeast cells well known in the art. Higher eukaryotic hosts mainly include mammalian cell lines known in the art and include many immortalized cell lines available from the ATCC, including HeLa cells, Chinese hamster  
5 ovary (CHO) cells, Baby hamster kidney (BHK) cells, PK15, RK13 and a number of other cell lines.

The present invention relates particularly to a recombinant E1 and/or E2 and/or E1/E2 protein expressed by a host cell as defined above containing a recombinant vector as defined above. These recombinant proteins are particularly purified according to the method  
10 of the present invention.

A preferred method for isolating or purifying HCV envelope proteins as defined above is further characterized as comprising at least the following steps:

- growing a host cell as defined above transformed with a recombinant vector according to the present invention or with a known recombinant vector expressing E1  
15 and/or E2 and/or E1/E2 HCV envelope proteins in a suitable culture medium,
- causing expression of said vector sequence as defined above under suitable conditions, and,
- lysing said transformed host cells, preferably in the presence of a SH group blocking agent, such as N-ethylmaleimide (NEM), and possibly a suitable detergent, preferably  
20 Empigen-BB,
- recovering said HCV envelope protein by affinity purification such as by means of lectin-chromatography or immunoaffinity chromatography using anti-E1 and/or anti-E2 specific monoclonal antibodies, with said lectin being preferably lentil-lectin or GNA, followed by,
- 25 - incubation of the eluate of the previous step with a disulphide bond cleavage means, such as DTT, preferably followed by incubation with an SH group blocking agent, such as NEM or Biotin-NEM, and,
- isolating the HCV single or specific oligomeric E1 and/or E2 and/or E1/E2 proteins such as by means of gel filtration and possibly also by a subsequent  $\text{Ni}^{2+}$ -IMAC  
30 chromatography followed by a desalting step.

As a result of the above-mentioned process, E1 and/or E2 and/or E1/E2 proteins may be produced in a form which elute differently from the large aggregates containing vector-derived components and/or cell components in the void volume of the gel filtration column or

the IMAC column as illustrated in the Examples section. The disulphide bridge cleavage step advantageously also eliminates the false reactivity due to the presence of host and/or expression-system-derived proteins. The presence of NEM and a suitable detergent during lysis of the cells may already partly or even completely prevent the aggregation between the HCV envelope proteins and contaminants.

Ni<sup>2+</sup>-IMAC chromatography followed by a desalting step is preferably used for constructs bearing a (His)<sub>6</sub> as described by Janknecht et al., 1991, and Hochuli et al., 1988.

The present invention also relates to a method for producing monoclonal antibodies in small animals such as mice or rats, as well as a method for screening and isolating human B-cells that recognize anti-HCV antibodies, using the HCV single or specific oligomeric envelope proteins of the present invention.

The present invention further relates to a composition comprising at least one of the following E1 peptides as listed in Table 3 and described elsewhere herein:

E1-31 (SEQ ID NO:56) spanning amino acids 181 to 200 of the Core/E1 V1 region,

E1-33 (SEQ ID NO:57) spanning amino acids 193 to 212 of the E1 region,

E1-35 (SEQ ID NO:58) spanning amino acids 205 to 224 of the E1 V2 region (epitope B),

E1-35A (SEQ ID NO:59) spanning amino acids 208 to 227 of the E1 V2 region (epitope B),

1bE1 (SEQ ID NO:53) spanning amino acids 192 to 228 of E1 regions (V1, C1, and V2 regions (containing epitope B)),

E1-51 (SEQ ID NO:66) spanning amino acids 301 to 320 of the E1 region,

E1-53 (SEQ ID NO:67) spanning amino acids 313 to 332 of the E1 C4 region (epitope A),

E1-55 (SEQ ID NO:68) spanning amino acids 325 to 344 of the E1 region,

The present invention also relates to a composition comprising at least one of the following E2 peptides as listed in Table 3:

Env 67 or E2-67 (SEQ ID NO:72) spanning amino acid positions 397 to 416 of the E2 region (epitope A, recognized by monoclonal antibody 2F10H10, see Figure 19),

Env 69 or E2-69 (SEQ ID NO:73) spanning amino acid positions 409 to 428 of the E2 region (epitope A),

Env 23 or E2-23 (SEQ ID NO:86) spanning positions 583 to 602 of the E2 region (epitope E),

Env 25 or E2-25 (SEQ ID NO:87) spanning positions 595 to 614 of the E2 region (epitope E),

Env 27 or E2-27 (SEQ ID NO:88) spanning positions 607 to 626 of the E2 region (epitope E),

5 Env 17B or E2-17B (SEQ ID NO:83) spanning positions 547 to 566 of the E2 region (epitope D),

Env 13B or E2-13B (SEQ ID NO:82) spanning positions 523 to 542 of the E2 region (epitope C; recognized by monoclonal antibody 16A6E7, see Figure 19).

10 The present invention also relates to a composition comprising at least one of the following E2 conformational epitopes:

epitope F recognized by monoclonal antibodies 15C8C1, 12D11F1 and 8G10D1H9,

epitope G recognized by monoclonal antibody 9G3E6,

epitope H (or C) recognized by monoclonal antibody 10D3C4 and 4H6B2, or,

epitope I recognized by monoclonal antibody 17F2C2.

15 The present invention also relates to an E1 or E2 specific antibody raised upon immunization with a peptide or protein composition, with said antibody being specifically reactive with any of the polypeptides or peptides as defined above, and with said antibody being preferably a monoclonal antibody.

20 The present invention also relates to an E1 or E2 specific antibody screened from a variable chain library in plasmids or phages or from a population of human B-cells by means of a process known in the art, with said antibody being reactive with any of the polypeptides or peptides as defined above, and with said antibody being preferably a monoclonal antibody.

25 The E1 or E2 specific monoclonal antibodies of the invention can be produced by any hybridoma liable to be formed according to classical methods from splenic cells of an animal, particularly from a mouse or rat, immunized against the HCV polypeptides or peptides according to the invention, as defined above on the one hand, and of cells of a myeloma cell line on the other hand, and to be selected by the ability of the hybridoma to produce the monoclonal antibodies recognizing the polypeptides which has been initially used for the immunization of the animals.

30 The antibodies involved in the invention can be labelled by an appropriate label of the enzymatic, fluorescent, or radioactive type.

The monoclonal antibodies according to this preferred embodiment of the invention may be humanized versions of mouse monoclonal antibodies made by means of recombinant



DNA technology, departing from parts of mouse and/or human genomic DNA sequences coding for H and L chains from cDNA or genomic clones coding for H and L chains.

Alternatively the monoclonal antibodies according to this preferred embodiment of the invention may be human monoclonal antibodies. These antibodies according to the present embodiment of the invention can also be derived from human peripheral blood lymphocytes of patients infected with HCV, or vaccinated against HCV. Such human monoclonal antibodies are prepared, for instance, by means of human peripheral blood lymphocytes (PBL) repopulation of severe combined immune deficiency (SCID) mice (for recent review, see Duchosal et al., 1992).

The invention also relates to the use of the proteins or peptides of the invention, for the selection of recombinant antibodies by the process of repertoire cloning (Persson et al., 1991).

Antibodies directed to peptides or single or specific oligomeric envelope proteins derived from a certain genotype may be used as a medicament, more particularly for incorporation into an immunoassay for the detection of HCV genotypes (for detecting the presence of HCV E1 or E2 antigen), for prognosing/monitoring of HCV disease, or as therapeutic agents.

Alternatively, the present invention also relates to the use of any of the above-specified E1 or E2 specific monoclonal antibodies for the preparation of an immunoassay kit for detecting the presence of E1 or E2 antigen in a biological sample, for the preparation of a kit for prognosing/monitoring of HCV disease or for the preparation of a HCV medicament.

The present invention also relates to the a method for *in vitro* diagnosis or detection of HCV antigen present in a biological sample, comprising at least the following steps :

- (i) contacting said biological sample with any of the E1 and/or E2 specific monoclonal antibodies as defined above, preferably in an immobilized form under appropriate conditions, which allow the formation of an immune complex,
- (ii) removing unbound components,
- (iii) incubating the immune complexes formed with heterologous antibodies, which specifically bind to the antibodies present in the sample to be analyzed, with said heterologous antibodies having conjugated to a detectable label under appropriate conditions,
- (iv) detecting the presence of said immune complexes visually or mechanically

(e.g. by means of densitometry, fluorimetry, colorimetry).

The present invention also relates to a kit for in vitro diagnosis of HCV antigen present in a biological sample, comprising:

- at least one monoclonal antibody as defined above, with said antibody being preferentially immobilized on a solid substrate,
- a buffer or components necessary for producing the buffer enabling binding reaction between these antibodies and the HCV antigens present in the biological sample,
- a means for detecting the immune complexes formed in the preceding binding reaction,
- possibly also including an automated scanning and interpretation device for inferring the HCV antigens present in the sample from the observed binding pattern.

The present invention also relates to a composition comprising E1 and/or E2 and/or E1/E2 recombinant HCV proteins purified according to the method of the present invention or a composition comprising at least one peptides as specified above for use as a medicament.

The present invention more particularly relates to a composition comprising at least one of the above-specified envelope peptides or a recombinant envelope protein composition as defined above, for use as a vaccine for immunizing a mammal, preferably humans, against HCV, comprising administering a sufficient amount of the composition possibly accompanied by pharmaceutically acceptable adjuvant(s), to produce an immune response.

More particularly, the present invention relates to the use of any of the compositions as described here above for the preparation of a vaccine as described above.

Also, the present invention relates to a vaccine composition for immunizing a mammal, preferably humans, against HCV, comprising HCV single or specific oligomeric proteins or peptides derived from the E1 and/or the E2 region as described above.

Immunogenic compositions can be prepared according to methods known in the art. The present compositions comprise an immunogenic amount of a recombinant E1 and/or E2 and/or E1/E2 single or specific oligomeric proteins as defined above or E1 or E2 peptides as defined above, usually combined with a pharmaceutically acceptable carrier, preferably further comprising an adjuvant.

The single or specific oligomeric envelope proteins of the present invention, either E1

and/or E2 and/or E1/E2, are expected to provide a particularly useful vaccine antigen, since the formation of antibodies to either E1 or E2 may be more desirable than to the other envelope protein, and since the E2 protein is cross-reactive between HCV types and the E1 protein is type-specific. Cocktails including type 1 E2 protein and E1 proteins derived from several genotypes may be particularly advantageous. Cocktails containing a molar excess of E1 versus E2 or E2 versus E1 may also be particularly useful. Immunogenic compositions may be administered to animals to induce production of antibodies, either to provide a source of antibodies or to induce protective immunity in the animal.

Pharmaceutically acceptable carriers include any carrier that does not itself induce the production of antibodies harmful to the individual receiving the composition. Suitable carriers are typically large, slowly metabolized macromolecules such as proteins, polysaccharides, polylactic acids, polyglycolic acids, polymeric amino acids, amino acid copolymers; and inactive virus particles. Such carriers are well known to those of ordinary skill in the art.

Preferred adjuvants to enhance effectiveness of the composition include, but are not limited to : aluminum hydroxide (alum), N-acetyl-muramyl-L-threonyl-D-isoglutamine (thr-MDP) as found in U.S. Patent No. 4,606,918, N-acetyl-normuramyl-L-alanyl-D-isoglutamine (nor-MDP), N-acetylmuramyl-L-alanyl-D-isoglutaminyl-L-alanine-2-(1'-2'-dipalmitoyl-sn-glycero-3-hydroxyphosphoryloxy)-ethylamine (MTP-PE) and RIBI, which contains three components extracted from bacteria, monophosphoryl lipid A, trehalose dimycolate, and cell wall skeleton (MPL+TDM+CWS) in a 2% squalene/Tween 80 emulsion. Any of the 3 components MPL, TDM or CWS may also be used alone or combined 2 by 2. Additionally, adjuvants such as Stimulon (Cambridge Bioscience, Worcester, MA) or SAF-1 (Syntex) may be used. Further, Complete Freund's Adjuvant (CFA) and Incomplete Freund's Adjuvant (IFA) may be used for non-human applications and research purposes.

The immunogenic compositions typically will contain pharmaceutically acceptable vehicles, such as water, saline, glycerol, ethanol, etc. Additionally, auxiliary substances, such as wetting or emulsifying agents, pH buffering substances, preservatives, and the like, may be included in such vehicles.

Typically, the immunogenic compositions are prepared as injectables, either as liquid solutions or suspensions; solid forms suitable for solution in, or suspension in, liquid vehicles prior to injection may also be prepared. The preparation also may be emulsified or encapsulated in liposomes for enhanced adjuvant effect. The E1 and E2 proteins may also be

incorporated into Immune Stimulating Complexes together with saponins, for example Quil A (ISCOMS).

Immunogenic compositions used as vaccines comprise a 'sufficient amount' or 'an immunologically effective amount' of the envelope proteins of the present invention, as well  
5 as any other of the above mentioned components, as needed. 'Immunologically effective amount', means that the administration of that amount to an individual, either in a single dose or as part of a series, is effective for treatment, as defined above. This amount varies depending upon the health and physical condition of the individual to be treated, the taxonomic group of individual to be treated (e.g. nonhuman primate, primate, etc.), the  
10 capacity of the individual's immune system to synthesize antibodies, the degree of protection desired, the formulation of the vaccine, the treating doctor's assessment of the medical situation, the strain of infecting HCV, and other relevant factors. It is expected that the amount will fall in a relatively broad range that can be determined through routine trials. Usually, the amount will vary from 0.01 to 1000 µg/dose, more particularly from 0.1 to 100  
15 µg/dose.

The single or specific oligomeric envelope proteins may also serve as vaccine carriers to present homologous (e.g. T cell epitopes or B cell epitopes from the core, NS2, NS3, NS4 or NS5 regions) or heterologous (non-HCV) haptens, in the same manner as Hepatitis B surface antigen (see European Patent Application 174,444). In this use, envelope proteins  
20 provide an immunogenic carrier capable of stimulating an immune response to haptens or antigens conjugated to the aggregate. The antigen may be conjugated either by conventional chemical methods, or may be cloned into the gene encoding E1 and/or E2 at a location corresponding to a hydrophilic region of the protein. Such hydrophilic regions include the V1 region (encompassing amino acid positions 191 to 202), the V2 region (encompassing  
25 amino acid positions 213 to 223), the V3 region (encompassing amino acid positions 230 to 242), the V4 region (encompassing amino acid positions 230 to 242), the V5 region (encompassing amino acid positions 294 to 303) and the V6 region (encompassing amino acid positions 329 to 336). Another useful location for insertion of haptens is the hydrophobic region (encompassing approximately amino acid positions 264 to 293). It is  
30 shown in the present invention that this region can be deleted without affecting the reactivity of the deleted E1 protein with antisera. Therefore, haptens may be inserted at the site of the deletion.

The immunogenic compositions are conventionally administered parenterally,

typically by injection, for example, subcutaneously or intramuscularly. Additional formulations suitable for other methods of administration include oral formulations and suppositories. Dosage treatment may be a single dose schedule or a multiple dose schedule. The vaccine may be administered in conjunction with other immunoregulatory agents.

5       The present invention also relates to a composition comprising peptides or polypeptides as described above, for *in vitro* detection of HCV antibodies present in a biological sample.

10       The present invention also relates to the use of a composition as described above for the preparation of an immunoassay kit for detecting HCV antibodies present in a biological sample.

      The present invention also relates to a method for *in vitro* diagnosis of HCV antibodies present in a biological sample, comprising at least the following steps :

- 15       (i)     contacting said biological sample with a composition comprising any of the envelope peptide or proteins as defined above, preferably in an immobilized form under appropriate conditions which allow the formation of an immune complex, wherein said peptide or protein can be a biotinylated peptide or protein which is covalently bound to a solid substrate by means of streptavidin or avidin complexes,
- 20       (ii)    removing unbound components,
- (iii)   incubating the immune complexes formed with heterologous antibodies, with said heterologous antibodies having conjugated to a detectable label under appropriate conditions,
- (iv)    detecting the presence of said immune complexes visually or mechanically (e.g. by means of densitometry, fluorimetry, colorimetry).

25       Alternatively, the present invention also relates to competition immunoassay formats in which recombinantly produced purified single or specific oligomeric protein E1 and/or E2 and/or E1/E2 proteins as disclosed above are used in combination with E1 and/or E2 peptides in order to compete for HCV antibodies present in a biological sample.

30       The present invention also relates to a kit for determining the presence of HCV antibodies, in a biological sample, comprising :

- at least one peptide or protein composition as defined above, possibly in combination with other polypeptides or peptides from HCV or other types of HCV, with said peptides or proteins being preferentially immobilized on a

solid substrate, more preferably on different microwells of the same ELISA plate, and even more preferentially on one and the same membrane strip,

- a buffer or components necessary for producing the buffer enabling binding reaction between these polypeptides or peptides and the antibodies against HCV present in the biological sample,
- means for detecting the immune complexes formed in the preceding binding reaction,
- possibly also including an automated scanning and interpretation device for inferring the HCV genotypes present in the sample from the observed binding pattern.

The immunoassay methods according to the present invention utilize single or specific oligomeric antigens from the E1 and/or E2 domains that maintain linear (in case of peptides) and conformational epitopes (single or specific oligomeric proteins) recognized by antibodies in the sera from individuals infected with HCV. It is within the scope of the invention to use for instance single or specific oligomeric antigens, dimeric antigens, as well as combinations of single or specific oligomeric antigens. The HCV E1 and E2 antigens of the present invention may be employed in virtually any assay format that employs a known antigen to detect antibodies. Of course, a format that denatures the HCV conformational epitope should be avoided or adapted. A common feature of all of these assays is that the antigen is contacted with the body component suspected of containing HCV antibodies under conditions that permit the antigen to bind to any such antibody present in the component. Such conditions will typically be physiologic temperature, pH and ionic strength using an excess of antigen. The incubation of the antigen with the specimen is followed by detection of immune complexes comprised of the antigen.

Design of the immunoassays is subject to a great deal of variation, and many formats are known in the art. Protocols may, for example, use solid supports, or immunoprecipitation. Most assays involve the use of labeled antibody or polypeptide; the labels may be, for example, enzymatic, fluorescent, chemiluminescent, radioactive, or dye molecules. Assays which amplify the signals from the immune complex are also known; examples of which are assays which utilize biotin and avidin or streptavidin, and enzyme-labeled and mediated immunoassays, such as ELISA assays.

The immunoassay may be, without limitation, in a heterogeneous or in a homogeneous format, and of a standard or competitive type. In a heterogeneous format, the

polypeptide is typically bound to a solid matrix or support to facilitate separation of the sample from the polypeptide after incubation. Examples of solid supports that can be used are nitrocellulose (e.g., in membrane or microtiter well form), polyvinyl chloride (e.g., in sheets or microtiter wells), polystyrene latex (e.g., in beads or microtiter plates, 5 polyvinylidene fluoride (known as Immunolon™), diazotized paper, nylon membranes, activated beads, and Protein A beads. For example, Dynatech Immunolon™ 1 or Immunolon™ 2 microtiter plates or 0.25 inch polystyrene beads (Precision Plastic Ball) can be used in the heterogeneous format. The solid support containing the antigenic polypeptides is typically washed after separating it from the test sample, and prior to detection of bound 10 antibodies. Both standard and competitive formats are known in the art.

In a homogeneous format, the test sample is incubated with the combination of antigens in solution. For example, it may be under conditions that will precipitate any antigen-antibody complexes which are formed. Both standard and competitive formats for these assays are known in the art.

15 In a standard format, the amount of HCV antibodies in the antibody-antigen complexes is directly monitored. This may be accomplished by determining whether labeled anti-xenogeneic (e.g. anti-human) antibodies which recognize an epitope on anti-HCV antibodies will bind due to complex formation. In a competitive format, the amount of HCV antibodies in the sample is deduced by monitoring the competitive effect on the binding of a 20 known amount of labeled antibody (or other competing ligand) in the complex.

Complexes formed comprising anti-HCV antibody (or in the case of competitive assays, the amount of competing antibody) are detected by any of a number of known techniques, depending on the format. For example, unlabeled HCV antibodies in the complex may be detected using a conjugate of anti-xenogeneic Ig complexed with a label (e.g. an 25 enzyme label).

In an immunoprecipitation or agglutination assay format the reaction between the HCV antigens and the antibody forms a network that precipitates from the solution or suspension and forms a visible layer or film of precipitate. If no anti-HCV antibody is present in the test specimen, no visible precipitate is formed.

30 There currently exist three specific types of particle agglutination (PA) assays. These assays are used for the detection of antibodies to various antigens when coated to a support. One type of this assay is the hemagglutination assay using red blood cells (RBCs) that are sensitized by passively adsorbing antigen (or antibody) to the RBC. The addition of specific

antigen antibodies present in the body component, if any, causes the RBCs coated with the purified antigen to agglutinate.

To eliminate potential non-specific reactions in the hemagglutination assay, two artificial carriers may be used instead of RBC in the PA. The most common of these are latex particles. However, gelatin particles may also be used. The assays utilizing either of these carriers are based on passive agglutination of the particles coated with purified antigens.

The HCV single or specific oligomeric E1 and/or E2 and/or E1/E2 antigens of the present invention comprised of conformational epitopes will typically be packaged in the form of a kit for use in these immunoassays. The kit will normally contain in separate containers the native HCV antigen, control antibody formulations (positive and/or negative), labeled antibody when the assay format requires the same and signal generating reagents (e.g. enzyme substrate) if the label does not generate a signal directly. The native HCV antigen may be already bound to a solid matrix or separate with reagents for binding it to the matrix. Instructions (e.g. written, tape, CD-ROM, etc.) for carrying out the assay usually will be included in the kit.

Immunoassays that utilize the native HCV antigen are useful in screening blood for the preparation of a supply from which potentially infective HCV is lacking. The method for the preparation of the blood supply comprises the following steps. Reacting a body component, preferably blood or a blood component, from the individual donating blood with HCV E1 and/or E2 proteins of the present invention to allow an immunological reaction between HCV antibodies, if any, and the HCV antigen. Detecting whether anti-HCV antibody - HCV antigen complexes are formed as a result of the reacting. Blood contributed to the blood supply is from donors that do not exhibit antibodies to the native HCV antigens, E1 or E2.

In cases of a positive reactivity to the HCV antigen, it is preferable to repeat the immunoassay to lessen the possibility of false positives. For example, in the large scale screening of blood for the production of blood products (e.g. blood transfusion, plasma, Factor VIII, immunoglobulin, etc.) 'screening' tests are typically formatted to increase sensitivity (to insure no contaminated blood passes) at the expense of specificity; i.e. the false-positive rate is increased. Thus, it is typical to only defer for further testing those donors who are 'repeatedly reactive'; i.e. positive in two or more runs of the immunoassay on the donated sample. However, for confirmation of HCV-positivity, the 'confirmation' tests are typically formatted to increase specificity (to insure that no false-positive samples are



confirmed) at the expense of sensitivity. Therefore the purification method described in the present invention for E1 and E2 will be very advantageous for including single or specific oligomeric envelope proteins into HCV diagnostic assays.

The solid phase selected can include polymeric or glass beads, nitrocellulose, 5 microparticles, microwells of a reaction tray, test tubes and magnetic beads. The signal generating compound can include an enzyme, a luminescent compound, a chromogen, a radioactive element and a chemiluminescent compound. Examples of enzymes include alkaline phosphatase, horseradish peroxidase and beta-galactosidase. Examples of enhancer compounds include biotin, anti-biotin and avidin. Examples of enhancer compounds binding 10 members include biotin, anti-biotin and avidin. In order to block the effects of rheumatoid factor-like substances, the test sample is subjected to conditions sufficient to block the effect of rheumatoid factor-like substances. These conditions comprise contacting the test sample with a quantity of anti-human IgG to form a mixture, and incubating the mixture for a time and under conditions sufficient to form a reaction mixture product substantially free of 15 rheumatoid factor-like substance.

The present invention further contemplates the use of E1 proteins, or parts thereof, more particularly HCV single or specific oligomeric E1 proteins as defined above, for *in vitro* monitoring HCV disease or prognosing the response to treatment (for instance with Interferon) of patients suffering from HCV infection comprising:

- 20 - incubating a biological sample from a patient with hepatitis C infection with an E1 protein or a suitable part thereof under conditions allowing the formation of an immunological complex,
- removing unbound components,
- calculating the anti-E1 titers present in said sample (for example at the start of 25 and/or during the course of (interferon) therapy),
- monitoring the natural course of HCV disease, or prognosing the response to treatment of said patient on the basis of the amount anti-E1 titers found in said sample at the start of treatment and/or during the course of treatment.

Patients who show a decrease of 2, 3, 4, 5, 7, 10, 15, or preferably more than 20 times 30 of the initial anti-E1 titers could be concluded to be long-term, sustained responders to HCV therapy, more particularly to interferon therapy. It is illustrated in the Examples section, that an anti-E1 assay may be very useful for prognosing long-term response to IFN treatment, or to treatment of Hepatitis C virus disease in general.

More particularly the following E1 peptides as listed in Table 3 were found to be useful for *in vitro* monitoring HCV disease or prognosing the response to interferon treatment of patients suffering from HCV infection:

- 5 E1-31 (SEQ ID NO:56) spanning amino acids 181 to 200 of the Core/E1 V1 region,
- E1-33 (SEQ ID NO:57) spanning amino acids 193 to 212 of the E1 region,
- E1-35 (SEQ ID NO:58) spanning amino acids 205 to 224 of the E1 V2 region (epitope B),
- E1-35A (SEQ ID NO:59) spanning amino acids 208 to 227 of the E1 V2 region (epitope B),
- 10 1bE1 (SEQ ID NO:53) spanning amino acids 192 to 228 of E1 regions (V1, C1, and V2 regions (containing epitope B)),
- E1-51 (SEQ ID NO:66) spanning amino acids 301 to 320 of the E1 region,
- E1-53 (SEQ ID NO:67) spanning amino acids 313 to 332 of the E1 C4 region (epitope A),
- 15 E1-55 (SEQ ID NO:68) spanning amino acids 325 to 344 of the E1 region.

It is to be understood that smaller fragments of the above-mentioned peptides also fall within the scope of the present invention. Said smaller fragments can be easily prepared by chemical synthesis and can be tested for their ability to be used in an assay as detailed above and in the Examples section.

- 20 The present invention also relates to a kit for monitoring HCV disease or prognosing the response to treatment (for instance to interferon) of patients suffering from HCV infection comprising:

- at least one E1 protein or E1 peptide, more particularly an E1 protein or E1 peptide as defined above,
- 25 - a buffer or components necessary for producing the buffer enabling the binding reaction between these proteins or peptides and the anti-E1 antibodies present in a biological sample,
- means for detecting the immune complexes formed in the preceding binding reaction,
- 30 - possibly also an automated scanning and interpretation device for inferring a decrease of anti-E1 titers during the progression of treatment.

It is to be understood that also E2 protein and peptides according to the present invention can be used to a certain degree to monitor/prognose HCV treatment as indicated

above for the E1 proteins or peptides because also the anti-E2 levels decrease in comparison to antibodies to the other HCV antigens. It is to be understood, however, that it might be possible to determine certain epitopes in the E2 region which would also be suited for use in an test for monitoring/prognosing HCV disease.

5           The present invention also relates to a serotyping assay for detecting one or more serological types of HCV present in a biological sample, more particularly for detecting antibodies of the different types of HCV to be detected combined in one assay format, comprising at least the following steps :

- 10           (i)     contacting the biological sample to be analyzed for the presence of HCV antibodies of one or more serological types, with at least one of the E1 and/or E2 and/or E1/E2 protein compositions or at least one of the E1 or E2 peptide compositions as defined above, preferentially in an immobilized form under appropriate conditions which allow the formation of an immune complex,
- (ii)     removing unbound components,
- 15           (iii)    incubating the immune complexes formed with heterologous antibodies, with said heterologous antibodies being conjugated to a detectable label under appropriate conditions,
- (iv)     detecting the presence of said immune complexes visually or mechanically (e.g. by means of densitometry, fluorimetry, colorimetry) and inferring the
- 20           presence of one or more HCV serological types present from the observed binding pattern.

It is to be understood that the compositions of proteins or peptides used in this method are recombinantly expressed type-specific envelope proteins or type-specific peptides.

25           The present invention further relates to a kit for serotyping one or more serological types of HCV present in a biological sample, more particularly for detecting the antibodies to these serological types of HCV comprising:

- at least one E1 and/or E2 and/or E1/E2 protein or E1 or E2 peptide, as defined above,
- 30           -           a buffer or components necessary for producing the buffer enabling the binding reaction between these proteins or peptides and the anti-E1-antibodies present in a biological sample,
- means for detecting the immune complexes formed in the preceding binding

reaction,

- possibly also an automated scanning and interpretation device for detecting the presence of one or more serological types present from the observed binding pattern.

5       The present invention also relates to the use of a peptide or protein composition as defined above, for immobilization on a solid substrate and incorporation into a reversed phase hybridization assay, preferably for immobilization as parallel lines onto a solid support such as a membrane strip, for determining the presence or the genotype of HCV according to a method as defined above. Combination with other type-specific antigens from other HCV  
10 polyprotein regions also lies within the scope of the present invention.

      The present invention provides a method for purifying recombinant HCV single or specific oligomeric envelope proteins selected from E1 and/or E2 and/or E1/E2 proteins which have been produced by a recombinant process comprising contacting said proteins with a disulphide bond cleavage or reducing agent. The contacting of the method of the  
15 invention may be carried out under partial cleavage or reducing conditions. Preferably, the disulphide bond cleavage agent is dithiothreitol (DTT), preferably in a concentration range of 0.1 to 50 mM, preferably 0.1 to 20 mM, more preferably 0.5 to 10 mM. Alternatively, the disulphide bond cleavage agent may be a detergent, such as Empigen-BB (which is a mixture containing N-Dodecyl-N,N-dimethylglycine as a major component), preferably at  
20 a concentration of 1 to 10%, more preferably at a concentration of 3.5%. Mixtures of detergents, disulphide bond cleavage agents and/or reducing agents may also be used. In one embodiment, disulphide bond reformation is prevented with an SH group blocking agent, such as N-ethylmaleimide (NEM) or a derivative thereof. In a preferred embodiment, the disulphide bond reformation is blocked by use of low pH conditions.

25       The present invention further provides a method as described herein, further involving the following steps: lysing recombinant E1 and/or E2 and/or E1/E2 expressing host cells, optionally in the presence of an SH blocking agent such as N-ethylmaleimide (NEM); recovering said HCV envelope proteins by affinity purification such as by means of lectin-chromatography, such as lentil-lectin chromatography, or by means of  
30 immunoaffinity using anti-E1 and/or anti-E2 specific monoclonal antibodies; reducing or cleaving of the disulfide bonds with a disulphide bond cleaving agent, such as DTT, preferably also in the presence of an SH blocking agent, such as NEM or Biotin-NEM; and, recovering the reduced E1 and/or E2 and/or E1/E2 envelope proteins by gelfiltration

and optionally additionally by a subsequent  $\text{Ni}^{2+}$ -IMAC chromatography and desalting step.

The present invention provides a composition containing substantially isolated and/or purified, and/or isolated and/or purified recombinant HCV single or specific  
5 oligomeric recombinant envelope proteins selected from E1 and/or E2 and/or E1/E2, which have preferably been isolated from the methods described herein. In a preferred embodiment, the recombinant HCV envelope proteins of the invention have been expressed in recombinant mammalian cells, such as vaccinia, recombinant yeast cells.

The present invention provides a recombinant vector containing a vector sequence,  
10 a prokaryotic, eukaryotic or viral promoter sequence and a nucleotide sequence allowing the expression of a single or specific oligomeric E1 and/or E2 and/or E1/E2 protein, in operable combination. In one embodiment, the nucleotide sequence of the vector encodes a single HCV E1 protein starting in the region between amino acid positions 1 and 192 and ending in the region between amino acid positions 250 and 400, more particularly ending  
15 in the region between positions 250 and 341, even more preferably ending in the region between position 290 and 341. In another embodiment, the nucleotide sequence of the vector encodes a single HCV E1 protein starting in the region between amino acid positions 117 and 192 and ending in the region between amino acid positions 263 and 400, more particularly ending in the region between positions 250 and 326. In yet another  
20 embodiment, the nucleotide sequence of the vector encodes a single HCV E1 protein bearing a deletion of the first hydrophobic domain between positions 264 to 293, plus or minus 8 amino acids. In a further embodiment, the nucleotide sequence of the vector encodes a single HCV E2 protein starting in the region between amino acid positions 290 and 406 and ending in the region between amino acid positions 600 and 820, more  
25 particularly starting in the region between positions 322 and 406, even more preferably starting in the region between position 347 and 406 and most preferably starting in the region between positions 364 and 406; and preferably ending at any of amino acid positions 623, 650, 661, 673, 710, 715, 720, 746 or 809. The vector of the present invention, in one embodiment, contains a 5'-terminal ATG codon and a 3'-terminal stop  
30 codon operably linked to the nucleotide sequence. The vector further contains, in one embodiment, a nucleotide sequence further containing at a factor Xa cleavage site and/or 3 to 10, preferably 6, histidine codons added 3'-terminally to the coding region. The vector of the present invention optionally contains a nucleotide sequence wherein at least one of

the glycosylation sites present in the E1 or E2 proteins has been removed at the nucleic acid level.

The present invention provides a nucleic acid containing any one of SEQ ID N0s: 3, 5, 7, 9, 11, 13, 21, 23, 25, 27, 29, 31, 35, 37, 39, 41, 43, 45, 47 and 49, or parts thereof.

5 The vector of the invention may preferably contain a nucleotide sequence containing a nucleic acid containing any one of SEQ ID N0s: 3, 5, 7, 9, 11, 13, 21, 23, 25, 27, 29, 31, 35, 37, 39, 41, 43, 45, 47 and 49, or parts thereof.

The composition of the present invention further contains recombinant HCV envelope proteins which have been expressed or are the expression product of a vector  
10 described herein.

The present invention provides a host cell transformed with at least one recombinant vector as described herein, wherein the vector contains a nucleotide sequence encoding HCV E1 and/or E2 and/or E1/E2 protein as described herein in addition to a regulatory sequence operable in the host cell and capable of regulating expression of the  
15 HCV E1 and/or E2 and/or E1/E2 protein. Moreover, the present invention provides a recombinant E1 and/or E2 and/or E1/E2 protein expressed by a host cell of the invention.

The present invention further provides a method as described herein and containing the following steps: growing a host cell as described herein which has been transformed with a recombinant vector as described herein in a suitable culture medium; causing  
20 expression of the vector nucleotide sequence of the vector, as described herein under suitable conditions; lysing the transformed host cells, preferably in the presence of an SH group blocking agent, such as N-ethylmaleimide (NEM); recovering the HCV envelope protein by affinity purification by means of for instance lectin-chromatography or immunoaffinity chromatography using anti-E1 and/or anti-E2 specific monoclonal  
25 antibodies, with said lectin being preferably lentil-lectin, followed by, incubation of the eluate of the previous step with a disulphide bond cleavage agent, such as DTT, preferably also in the presence of an SH group blocking agent, such as NEM or Biotin-NEM; and, isolating the HCV single or specific oligomeric E1 and/or E2 and/or E1/E2 proteins by means of gelfiltration and possibly also by means of an additional Ni<sup>2+</sup>-IMAC  
30 chromatography and desalting step.

The present invention provides a composition containing at least one of the following E1 and/or E2 peptides:

E1-31 (SEQ ID NO:56) spanning amino acids 181 to 200 of the Core/E1

V1 region,

E1-33 (SEQ ID NO:57) spanning amino acids 193 to 212 of the E1 region,

E1-35 (SEQ ID NO:58) spanning amino acids 205 to 224 of the E1 V2 region (epitope B),

5 E1-35A (SEQ ID NO:59) spanning amino acids 208 to 227 of the E1 V2 region (epitope B),

1bE1 (SEQ ID NO:53) spanning amino acids 192 to 228 of E1 regions (V1, C1, and V2 regions (containing epitope B),

E1-51 (SEQ ID NO:66) spanning amino acids 301 to 320 of the E1 region,

10 E1-53 (SEQ ID NO:67) spanning amino acids 313 to 332 of the E1 C4 region (epitope A),

E1-55 (SEQ ID NO:68) spanning amino acids 325 to 344 of the E1 region.

Env 67 or E2-67 (SEQ ID NO:72) spanning amino acid positions 397 to 416 of the E2 region (epitope A),

15 Env 69 or E2-69 (SEQ ID NO:73) spanning amino acid positions 409 to 428 of the E2 region (epitope A),

Env 23 or E2-23 (SEQ ID NO:86) spanning positions 583 to 602 of the E2 region (epitope E),

20 Env 25 or E2-25 (SEQ ID NO:87) spanning positions 595 to 614 of the E2 region (epitope E),

Env 27 or E2-27 (SEQ ID NO:88) spanning positions 607 to 626 of the E2 region (epitope E),

Env 17B or E2-17B (SEQ ID NO:83) spanning positions 547 to 566 of the E2 region (epitope D),

25 Env 13B or E2-13B (SEQ ID NO:82) spanning positions 523 to 542 of the E2 region (epitope C).

The present invention provides a composition containing at least one of the following E2 conformational epitopes:

30 epitope F recognized by monoclonal antibodies 15C8C1, 12D11F1, and 8G10D1H9,

epitope G recognized by monoclonal antibody 9G3E6,

epitope H (or C) recognized by monoclonal antibodies 10D3C4 and 4H6B2,

epitope I recognized by monoclonal antibody 17F2C2.

The present invention provides an E1 and/or E2 specific monoclonal antibody raised upon immunization with a composition as described herein. The antibodies of the present invention may be used, for example, as a medicament, for incorporation into an immunoassay kit for detecting the presence of HCV E1 or E2 antigen, for prognosis/monitoring of disease or for HCV therapy. The present invention provides for the use of an E1 and/or E2 specific monoclonal antibody as described herein for the preparation of an immunoassay kit for detecting HCV E1 or E2 antigens, for the preparation of a kit for prognosing/monitoring of HCV disease or for the preparation of a HCV medicament.

10 The present invention provides a method for in vitro diagnosis of HCV antigen present in a biological sample, containing at least the following steps:

- (i) contacting said biological sample with an E1 and/or E2 specific monoclonal antibody as described herein, preferably in an immobilized form under appropriate conditions which allow the formation of an immune complex,
- 15 (ii) removing unbound components,
- (iii) incubating the immune complexes formed with heterologous antibodies, with the heterologous antibodies being conjugated to a detectable label under appropriate conditions,
- (iv) detecting the presence of the immune complexes visually or mechanically.

20 The present invention provides a kit for determining the presence of HCV antigens present in a biological sample, which includes at least the following: at least one E1 and/or E2 specific monoclonal antibody as described herein, preferably in an immobilized form on a solid substrate, a buffer or components necessary for producing the buffer enabling binding reaction between these antibodies and the HCV antigens present in a biological sample, and optionally a means for detecting the immune complexes formed in the preceding binding reaction.

The composition of the present invention may be provided in the form of a medicament.

30 The present invention provides a composition, as described herein for use as a vaccine for immunizing a mammal, preferably humans, against HCV, comprising administering an effective amount of said composition being optionally accompanied by pharmaceutically acceptable adjuvants, to produce an immune response.

The present invention provides a method of using the composition, as described



herein, for the preparation of a vaccine for immunizing a mammal, preferably humans, against HCV, comprising administering an effective amount of said composition, optionally accompanied by pharmaceutically acceptable adjuvants, to produce an immune response.

5       The present invention provides a vaccine composition for immunizing a mammal, preferably humans, against HCV, which contains an effective amount of a composition containing an E1 and/or E2 containing composition as described herein, optionally also accompanied by pharmaceutically acceptable adjuvants.

10       The composition of the present invention may be provided in a form for *in vitro* detection of HCV antibodies present in a biological sample. The present invention also provides a method of preparing an immunoassay kit for detecting HCV antibodies present in a biological sample and a method of detecting HCV antibodies present in a biological sample using the kit of the invention to diagnose HCV antibodies present in a biological sample. Such a method of the present invention includes at least the following steps:

15       (i) contacting said biological sample with a composition as described herein, preferably in an immobilized form under appropriate conditions which allow the formation of an immune complex with HCV antibodies present in the biological sample,

(ii) removing unbound components,

20       (iii) incubating the immune complexes formed with heterologous antibodies, with the heterologous antibodies being conjugated to a detectable label under appropriate conditions,

(iv) detecting the presence of the immune complexes visually or mechanically.

25       The present invention provides a kit for determining the presence of HCV antibodies present in a biological sample, containing: at least one peptide or protein composition as described herein, preferably in an immobilized form on a solid substrate; a buffer or components necessary for producing the buffer enabling binding reaction between these proteins or peptides and the antibodies against HCV present in the biological sample; and, optionally, a means for detecting the immune complexes formed in the preceding binding reaction.

30       The present invention provides a method of *in vitro* monitoring HCV disease or diagnosing the response of a patient suffering from HCV infection to treatment, preferably with interferon, the method including: incubating a biological sample from the patient with HCV infection with an E1 protein or a suitable part thereof under conditions allowing the

formation of an immunological complex; removing unbound components; calculating the anti-E1 titers present in the sample at the start of and during the course of treatment; monitoring the natural course of HCV disease, or diagnosing the response to treatment of the patient on the basis of the amount anti-E1 titers found in the sample at the start of  
5 treatment and/or during the course of treatment.

The present invention provides a kit for monitoring HCV disease or prognosing the response to treatment, particularly with interferon, of patients suffering from HCV infection, wherein the kit contains: at least one E1 protein or E1 peptide, more particularly an E1 protein or E1 peptide as described herein; a buffer or components necessary for  
10 producing the buffer enabling the binding reaction between these proteins or peptides and the anti-E1 antibodies present in a biological sample; and optionally, means for detecting the immune complexes formed in the preceding binding reaction, optionally, also an automated scanning and interpretation device for inferring a decrease of anti-E1 titers during the progression of treatment.

15 The present invention provides a serotyping assay for detecting one or more serological types of HCV present in a biological sample, more particularly for detecting antibodies of the different types of HCV to be detected combined in one assay format, including at least the following steps: (i) contacting the biological sample to be analyzed for the presence of HCV antibodies of one or more serological types, with at least one of  
20 the E1 and/or E2 and/or E1/E2 protein compositions as described herein or at least one of the E1 or E2 peptide compositions described herein, preferentially in an immobilized form under appropriate conditions which allow the formation of an immune complex; (ii) removing unbound components; (iii) incubating the immune complexes formed with heterologous antibodies, with the heterologous antibodies being conjugated to a detectable  
25 label under appropriate conditions; and optionally, (iv) detecting the presence of said immune complexes visually or mechanically (e.g. by means of densitometry, fluorimetry, colorimetry) and inferring the presence of one or more HCV serological types present from the observed binding pattern.

The present invention provides a kit for serotyping one or more serological types of  
30 HCV present in a biological sample, more particularly for detecting the antibodies to these serological types of HCV containing: at least one E1 and/or E2 and/or E1/E2 protein as described herein or an E1 or E2 peptide as described herein; a buffer or components necessary for producing the buffer enabling the binding reaction between these proteins or

peptides and the anti-E1 antibodies present in a biological sample; optionally, means for detecting the immune complexes formed in the preceding binding reaction, optionally, also an automated scanning and interpretation device for detecting the presence of one or more serological types present from the observed binding pattern.

5       The present invention provides a peptide or protein composition as described herein, for immobilization on a solid substrate and incorporation into a reversed phase hybridization assay, preferably for immobilization as parallel lines onto a solid support such as a membrane strip, for determining the presence or the genotype of HCV according to a method as described herein.

10       The present invention provides a therapeutic vaccine composition, such as a therapeutic HCV vaccine composition, containing or comprising a therapeutically effective amount of: a composition containing at least one purified recombinant HCV single or specific oligomeric recombinant envelope proteins selected from the group consisting of an E1 protein, an E2 protein, a part of said E1 and E2 proteins, an E1/E2 protein complex  
15       formed from purified HCV single or specific oligomeric recombinant E1 or E2 proteins or parts thereof; and optionally a pharmaceutically acceptable adjuvant. Another therapeutic vaccine composition, such as a therapeutic HCV vaccine composition, of the invention may be comprising a therapeutically effective amount of a combination of at least two purified HCV single or specific oligomeric recombinant envelope proteins selected from  
20       the group consisting of E1 proteins derived from different HCV genotypes or subtypes, E2 proteins derived from different HCV genotypes or subtypes, parts of said E1 and E2 proteins, and E1/E2 protein complexes formed from purified HCV single or specific oligomeric recombinant E1 or E2 proteins, or parts thereof, derived from different HCV genotypes or subtypes; and optionally a pharmaceutically acceptable adjuvant.

25       The HCV envelope proteins of the vaccine, or more particularly the HCV vaccine, of the present invention are optionally produced by recombinant mammalian cells, by recombinant yeast cells, or by or via a recombinant virus. The invention provides a therapeutic vaccine composition, such as a therapeutic HCV vaccine composition, containing or comprising a therapeutically effective amount of a composition comprising  
30       at least one of the following E1 and E2 peptides:

    E1-31 (SEQ ID NO:56) spanning amino acids 181 to 200 of the Core/E1 V1 region,

    E1-33 (SEQ ID NO:57) spanning amino acids 193 to 212 of the E1 region,

    E1-35 (SEQ ID NO:58) spanning amino acids 205 to 224 of the E1 V2 region (epitope

B),

E1-35A (SEQ ID NO:59) spanning amino acids 208 to 227 of the E1 V2 region (epitope B),

1bE1 (SEQ ID NO:53) spanning amino acids 192 to 228 of E1 regions V1, C1, and V2 regions (containing epitope B),

E1-51 (SEQ ID NO:66) spanning amino acids 301 to 320 of the E1 region,

E1-53 (SEQ ID NO:67) spanning amino acids 313 to 332 of the E1 C4 region (epitope A),

E1-55 (SEQ ID NO:68) spanning amino acids 325 to 344 of the E1 region,

Env 67 or E2-67 (SEQ ID NO:72) spanning amino acid positions 397 to 418 of the E2 region (epitope A),

Env 69 or E2-69 (SEQ ID NO:73) spanning amino acid positions 409 to 428 of the E2 region (epitope A),

Env 23 or E2-23 (SEQ ID NO:86) spanning positions 583 to 602 of the E2 region (epitope E),

Env 25 or E2-25 (SEQ ID NO:87) spanning positions 595 to 614 of the E2 region (epitope E),

Env 27 or E2-27 (SEQ ID NO:88) spanning positions 607 to 626 of the E2 region (epitope E),

Env 17B or E2-17B (SEQ ID NO:83) spanning positions 547 to 586 of the E2 region (epitope D),

Env 13B or E2-13B (SEQ ID NO:82) spanning positions 523 to 542 of the E2 region (epitope C),

IGP 1626 spanning positions 192-211 of the E1 region (SEQ ID NO:112),

IGP 1627 spanning positions 204-223 of the E1 region (SEQ ID NO:113),

IGP 1628 spanning positions 216-235 of the E1 region (SEQ ID NO:114),

IGP 1629 spanning positions 228-247 of the E1 region (SEQ ID NO:115),

IGP 1630 spanning positions 240-259 of the E1 region (SEQ ID NO:116),

IGP 1631 spanning positions 252-271 of the E1 region (SEQ ID NO:117),

IGP 1632 spanning positions 264-283 of the E1 region (SEQ ID NO:118),

IGP 1633 spanning positions 276-295 of the E1 region (SEQ ID NO:119),

IGP 1634 spanning positions 288-307 of the E1 region (SEQ ID NO:120),

IGP 1635 spanning positions 300-319 of the E1 region (SEQ ID NO:121) and

IGP 1636 spanning positions 312-331 of the E1 region (SEQ ID NO:122); and wherein said peptides may be of recombinant or synthetic origin, and optionally combined with a pharmaceutically acceptable adjuvant.

Any of the above mentioned therapeutic vaccine compositions may also be considered or regarded as therapeutic HCV vaccine compositions, or as therapeutic compositions or therapeutic HCV compositions, or as compositions or HCV compositions.

The present invention provides a method of treating a mammal, such as a human, infected with HCV comprising administering an effective amount of a composition as described herein, such as the above described vaccines or therapeutic compositions, and optionally, a pharmaceutically acceptable adjuvant. In one embodiment, the composition of the invention is administered in combination with or at a time in conjunction with antiviral therapy, either soon prior to or subsequent to or with administration of the composition of the invention. It will be clear that any of the compositions of the invention, e.g. a therapeutic HCV vaccine composition, can be used for treating a mammal chronically infected with HCV (a "chronic HCV-infected mammal").

The present invention provides a composition, such as a therapeutic HCV composition or a HCV composition, containing or comprising at least one purified recombinant HCV recombinant envelope proteins selected from the group consisting of an E1 protein and an E2 protein, and optionally an adjuvant. In a preferred embodiment, the composition contains at least one of the following E1 and E2 peptides:

E1-31 (SEQ ID NO:56) spanning amino acids 181 to 200 of the Core/E1 V1 region,

E1-33 (SEQ ID NO:57) spanning amino acids 193 to 212 of the E1 region,

E1-35 (SEQ ID NO:58) spanning amino acids 205 to 224 of the E1 V2 region (epitope B),

E1-35A (SEQ ID NO:59) spanning amino acids 208 to 227 of the E1 V2 region (epitope B),

1bE1 (SEQ ID NO:53) spanning amino acids 192 to 228 of E1 regions V1, C1, and V2 regions (containing epitope B),

E1-51 (SEQ ID NO:66) spanning amino acids 301 to 320 of the E1 region,

E1-53 (SEQ ID NO:67) spanning amino acids 313 to 332 of the E1 C4 region (epitope A),

E1-55 (SEQ ID NO:68) spanning amino acids 325 to 344 of the E1 region,

Env 67 or E2-67 (SEQ ID NO:72) spanning amino acid positions 397 to 418 of the E2

- region (epitope A),  
Env 69 or E2-69 (SEQ ID NO:73) spanning amino acid positions 409 to 428 of the E2 region (epitope A),  
Env 23 or E2-23 (SEQ ID NO:86) spanning positions 583 to 602 of the E2 region (epitope E),  
5 Env 25 or E2-25 (SEQ ID NO:87) spanning positions 595 to 614 of the E2 region (epitope E),  
Env 27 or E2-27 (SEQ ID NO:88) spanning positions 607 to 626 of the E2 region (epitope E),  
10 Env 17B or E2-17B (SEQ ID NO:83) spanning positions 547 to 586 of the E2 region (epitope D),  
Env 13B or E2-13B (SEQ ID NO:82) spanning positions 523 to 542 of the E2 region (epitope C),  
IGP 1626 spanning positions 192-211 of the E1 region (SEQ ID NO:112),  
15 IGP 1627 spanning positions 204-223 of the E1 region (SEQ ID NO:113),  
IGP 1628 spanning positions 216-235 of the E1 region (SEQ ID NO:114),  
IGP 1629 spanning positions 228-247 of the E1 region (SEQ ID NO:115),  
IGP 1630 spanning positions 240-259 of the E1 region (SEQ ID NO:116),  
IGP 1631 spanning positions 252-271 of the E1 region (SEQ ID NO:117),  
20 IGP 1632 spanning positions 264-283 of the E1 region (SEQ ID NO:118),  
IGP 1633 spanning positions 276-295 of the E1 region (SEQ ID NO:119),  
IGP 1634 spanning positions 288-307 of the E1 region (SEQ ID NO:120),  
IGP 1635 spanning positions 300-319 of the E1 region (SEQ ID NO:121) and  
IGP 1636 spanning positions 312-331 of the E1 region (SEQ ID NO:122),  
25 and wherein said peptides may be of recombinant or synthetic origin, and optionally combined with a pharmaceutically acceptable adjuvant.

Another composition, such as a therapeutic HCV composition or HCV composition, of the invention may be comprising a therapeutically effective amount of a combination of at least two purified HCV single or specific oligomeric recombinant  
30 envelope proteins selected from the group consisting of E1 proteins derived from different HCV genotypes or subtypes, E2 proteins derived from different HCV genotypes or subtypes, parts of said E1 and E2 proteins, and E1/E2 protein complexes formed from purified HCV single or specific oligomeric recombinant E1 or E2 proteins, or parts thereof,

derived from different HCV genotypes or subtypes; and optionally a pharmaceutically acceptable adjuvant.

The present invention provides a therapeutic composition or therapeutic vaccine composition or composition for inducing HCV-specific antibodies or for stimulating T-cell activity or for stimulating cytokine secretion or cytokine production. The present invention provides a therapeutic HCV composition or therapeutic HCV vaccine composition or HCV composition for inducing HCV-specific antibodies or for stimulating T-cell activity or for stimulating cytokine secretion or cytokine production.

The therapeutic composition according to the present invention, such as a therapeutic HCV vaccine composition or therapeutic HCV composition, may also be therapeutically effective in a HCV carrier infected with a HCV genotype different from the HCV genotype or HCV genotypes from which said E1, E2, or E1/E2 protein complexes are derived.

The recombinant HCV envelope proteins may be produced by recombinant mammalian cells, recombinant HCV envelope proteins are produced by recombinant yeast cells, or recombinant HCV envelope proteins are produced by or via a recombinant virus. The present invention provides a method of treating a mammal, such as a human, infected with HCV including administering an effective amount of a composition described herein, such as a therapeutic HCV vaccine composition, and, optionally, a pharmaceutically acceptable adjuvant. It will be clear that any of the compositions of the invention, such as a therapeutic HCV vaccine composition, can be used for treating a chronic HCV-infected mammal. The present invention provides a therapeutic composition for inducing HCV-specific antibodies, for stimulating T-cell activity or for stimulating cytokine secretion or cytokine production, said composition containing a therapeutic effective amount of a composition containing at least one purified recombinant HCV single or specific oligomeric recombinant envelope protein selected from the group consisting of an E1 protein and an E2 protein; and optionally a pharmaceutically acceptable adjuvant.

In particular, any of the compositions according to this invention (including the vaccine compositions and therapeutic compositions) may comprise recombinant HCV envelope proteins wherein the cysteines of said recombinant HCV envelope proteins are blocked, or may comprise E1 and/or E2 peptides wherein the cysteines of said E1 or E2 peptides are blocked.

In another embodiment of the invention, the compositions (including the vaccine

compositions and therapeutic compositions) according to the invention may comprise recombinant HCV envelope proteins which are added to said compositions as viral-like particles (VLPs).

In a further embodiment, a composition of the invention such as a therapeutic HCV vaccine composition may comprise as recombinant HCV E1 envelope protein an E1s protein. More particularly, said E1s protein is defined by SEQ ID NO:123.

Another aspect of the invention relates to an immunogenic composition, in particular a HCV immunogenic composition, comprising a recombinant virus allowing expression of at least one HCV recombinant envelope protein chosen from an E1 protein and/or an E2 protein, and parts of said E1 and E2 proteins; and, optionally, a pharmaceutically acceptable adjuvant.

In yet a further aspect, the invention is envisaging a vaccine composition such as a HCV vaccine composition comprising a recombinant virus allowing expression of at least one HCV recombinant envelope protein chosen from an E1 protein and/or an E2 protein, and parts of said E1 and E2 proteins; and, optionally, a pharmaceutically acceptable adjuvant.

In one embodiment, the above recombinant virus compositions may be effective against a HCV genotype or subtype different from the HCV genotype or subtype from which said E1 protein and/or E2 protein, or said parts thereof, are derived.

In another embodiment, the above recombinant virus compositions may be used for inducing HCV-specific antibodies or for stimulating T-cell activity or for stimulating cytokine secretion or cytokine production.

In another embodiment to the recombinant virus compositions, said recombinant virus is a vaccinia virus, a recombinant avipox virus or a recombinant Ankara Modified Virus.

Another aspect of the invention relates to a method of treating a mammal infected with HCV comprising administering an effective amount of a recombinant vaccine composition as described above.

The mammal in any of the above aspects of the invention may in particular be a human.

One further aspect of the present invention relates to a method to reduce liver disease in a chronic HCV-infected mammal or human comprising administering a therapeutic vaccine to said mammal or human.

In another aspect, a method to reduce liver fibrosis progression in a chronic HCV-



infected mammal or human comprising administering a therapeutic vaccine to said mammal or human is covered.

5 A further aspect relates to a method to reduce liver fibrosis in a chronic HCV-infected mammal or human comprising administering a therapeutic vaccine to said mammal or human.

A further aspect relates to a method to reduce liver steatosis in a chronic HCV-infected mammal or human comprising administering a therapeutic vaccine to said mammal or human.

10 Yet another aspect of the invention embodies a method to reduce liver disease by at least 2 points according to the overall Ishak score or Ishak activity score in a chronic HCV-infected mammal or human comprising administering a therapeutic vaccine to said mammal or human.

15 Another further aspect relates to a method to reduce liver disease or liver fibrosis by at least 1 point according to the Ishak fibrosis score in a chronic HCV-infected mammal or human comprising administering a therapeutic vaccine to said mammal or human.

A further aspect of the invention provides a method to reduce serum ALT levels in a chronic HCV-infected mammal or human comprising administering a therapeutic vaccine to said mammal or human.

20 Yet another aspect of the invention relates to a method to reduce anti-E1 and/or anti-E2 immunoreactivity in the liver of a chronic HCV-infected mammal or human comprising administering a therapeutic vaccine to said mammal or human.

25 Furthermore included in the invention is a method for treating a chronic HCV-infected mammal or human wherein said method is comprising administration of multiple doses of any of the compositions of the invention, such as a therapeutic HCV vaccine composition, to said mammal or human and wherein said multiple doses are administrated to said mammal or human separated by a specified time interval. Thus, said plurality of administrations of a composition of the invention to treat a chronic HCV-infected carrier may be separated, e.g., by a time-interval of 4 weeks or less. Thus, said time intervals could be 1 or 1.5 or 2 or 2.5 or 3 or 3.5 or 4 weeks, or could be 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 30 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27 or 28 days. In particular, said method of treating a chronic HCV-infected mammal or human comprises a plurality of administrations of a composition of the invention, such as a therapeutic HCV vaccine composition, to said mammal or human wherein said administrations are separated by a

time interval of 3 weeks. In a further embodiment, said plurality of administrations consists of a first series of at least 5 administrations followed by an administration-free period of at least 12 weeks followed by a second series of at least 3 administrations. In particular, a series of administrations may comprise at least 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, or  
5 more administrations. The administration-free period may be at least 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18 or more weeks.

It will furthermore be clear that any of the compositions of the invention, such as a therapeutic HCV vaccine composition, can be used for obtaining any of the effects aimed at in the various methods as described above. Said therapeutic vaccine as applied in the  
10 methods described above may be any composition of the current invention, such as a therapeutic HCV vaccine composition.

The use of the HCV E1/E2 proteins or parts thereof or E1/E2 peptides as outlined throughout the above description for the manufacture of a composition, HCV composition, vaccine, HCV vaccine, therapeutic vaccine or therapeutic HCV vaccine for use in any of  
15 the methods outlined above or used for obtaining any of the effects aimed at in the various methods outlined above is also envisaged by the current invention.

In any of these methods said therapeutic vaccine comprises at least one HCV antigen and, optionally, a pharmaceutically acceptable adjuvant such as alum. In one embodiment thereto said HCV antigen is an E1 or E2 antigen, or an immunogenic part of  
20 an E1 or E2 antigen. When referring to the HCV antigen as being an E1s antigen, the E1s antigen may be defined by SEQ ID NO:123.

With the term "liver disease" is meant in this context any abnormal liver condition caused by infection with the hepatitis C virus including inflammation, fibrosis, cirrhosis, necrosis, necro-inflammation and hepatocellular carcinoma.

25 With "steatosis" is meant a histological feature of lipid accumulation in the hepatocytes that is indicative of liver involvement in a wide variety of systemic disorders, toxic or drug-induced liver injury, as well as of various specific liver diseases, including hepatitis C infection, Wilson's disease, and galactosemia.

With "reducing liver disease" is meant any stabilization or reduction of the liver  
30 disease status. Liver disease can be determined, e.g., by the Knodell scoring system or the Knodell scoring system adapted by Ishak. A reduction of this score by two points is accepted as therapeutically beneficial effect in several studies (see, e.g., studies published after 1996 as indicated in Table 2 of Shiffman 1999).

With "reducing liver fibrosis progression" is meant any slowing down, halting or reverting of the normally expected progression of liver fibrosis. Liver fibrosis progression can be determined, e.g., by the Metavir scoring system. Normal expected progression of liver fibrosis according to this system was published to be an increase of the Metavir score of an untreated chronic HCV patient of approximately 0.133 per year (Poynard et al. 1997). Fibrosis is considered to include any form of fibrosis, e.g. as scored by the Metavir or Ishak system, including perisinusoidal fibrosis.

With the term "HCV antigen" is meant any HCV protein or fragment thereof comprising at least one T cell epitope or B cell epitope.

A further aspect of the current invention provides a method to predict changes in liver disease in a chronic HCV-infected mammal or human, said method comprising:

- (i) determining the level of serum anti-E1 antibody level prior to therapeutic vaccination with a HCV vaccine composition comprising an E1 antigen;
- (ii) determining the level of serum anti-E1 antibody level after therapeutic vaccination with a HCV vaccine composition comprising an E1 antigen;
- (iii) inferring the difference in level of serum anti-E1 antibody level determined in (i) and (ii) and, therefrom, predicting the change in liver disease.

It will be clear to the skilled artisan that a significantly high and positive difference in level of serum anti-E1 antibody, calculated as the value measured in (ii) minus the value measured in (i), would tip the balance in favor of a positive prediction, i.e., a prediction that the degree of liver disease will decrease. It will also be clear that a sustained significantly high and positive difference would add to the positive character of the prediction of decreased liver disease. When said difference is, however, not significantly high, is zero, is negative, or is significantly high and positive at one point in time but is not sustained, this would then tip the balance in favor of a negative prediction, i.e., a prediction that the degree of liver disease will remain unchanged or will increase. A sustained high level of serum anti-E1 antibody could be reached through additional immunizations either by administering a new series of immunizations after an administration free period or by repeating immunizations with a larger time interval, e.g. 6 weeks, after an initial priming series consisting of administrations with a short time interval, e.g. 3 weeks.

**Figure and Table legends**

- Figure 1 : Restriction map of plasmid pgpt ATA 18
- Figure 2 : Restriction map of plasmid pgs ATA 18
- 5 Figure 3 : Restriction map of plasmid pMS 66
- Figure 4 : Restriction map of plasmid pv HCV-11A
- Figure 5 : Anti-E1 levels in non-responders to IFN treatment
- Figure 6 : Anti-E1 levels in responders to IFN treatment
- Figure 7 : Anti-E1 levels in patients with complete response to IFN treatment
- 10 Figure 8 : Anti-E1 levels in incomplete responders to IFN treatment
- Figure 9 : Anti-E2 levels in non-responders to IFN treatment
- Figure 10 : Anti-E2 levels in responders to IFN treatment
- Figure 11 : Anti-E2 levels in incomplete responders to IFN treatment
- Figure 12 : Anti-E2 levels in complete responders to IFN treatment
- 15 Figure 13 : Human anti-E1 reactivity competed with peptides
- Figure 14 : Competition of reactivity of anti-E1 monoclonal antibodies with peptides
- Figure 15 : Anti-E1 (epitope 1) levels in non-responders to IFN treatment
- Figure 16 : Anti-E1 (epitope 1) levels in responders to IFN treatment
- Figure 17 : Anti-E1 (epitope 2) levels in non-responders to IFN treatment
- 20 Figure 18 : Anti-E1 (epitope 2) levels in responders to IFN treatment
- Figure 19 : Competition of reactivity of anti-E2 monoclonal antibodies with peptides
- Figure 20 : Human anti-E2 reactivity competed with peptides
- Figure 21 : Nucleic acid sequences of the present invention. The nucleic acid sequences encoding an E1 or E2 protein according to the present invention may be translated (SEQ ID NO 3 to 13, 21-31, 35 and 41-49 are translated in a reading frame starting from residue number 1, SEQ ID NO 37-39 are translated in a reading frame starting from residue number 2), into the amino acid sequences of the respective E1 or E2 proteins as shown in the sequence listing.
- 25
- 30 Figure 22 : ELISA results obtained from lentil lectin chromatography eluate fractions of 4 different E1 purifications of cell lysates infected with vvHCV39 (type 1b), vvHCV40 (type 1b), vvHCV62 (type 3a), and vvHCV63 (type 5a).
- Figure 23 : Elution profiles obtained from the lentil lectin chromatography of the 4

different E1 constructs on the basis of the values as shown in Figure 22.

Figure 24: ELISA results obtained from fractions obtained after gelfiltration chromatography of 4 different E1 purifications of cell lysates infected with vvHCV39 (type 1b), vvHCV40 (type 1b), vvHCV62 (type 3a), and vvHCV63 (type 5a).

Figure 25: Profiles obtained from purifications of E1 proteins of type 1b (1), type 3a (2), and type 5a (3) (from RK13 cells infected with vvHCV39, vvHCV62, and vvHCV63, respectively; purified on lentil lectin and reduced as in example 5.2 - 5.3) and a standard (4). The peaks indicated with '1', '2', and '3', represent pure E1 protein peaks (see Figure 24, E1 reactivity mainly in fractions 26 to 30).

Figure 26: Silver staining of an SDS-PAGE as described in example 4 of a raw lysate of E1 vvHCV40 (type 1b) (lane 1), pool 1 of the gelfiltration of vvHCV40 representing fractions 10 to 17 as shown in Figure 25 (lane 2), pool 2 of the gelfiltration of vvHCV40 representing fractions 18 to 25 as shown in Figure 25 (lane 3), and E1 pool (fractions 26 to 30) (lane 4).

Figure 27: Streptavidine-alkaline phosphatase blot of the fractions of the gelfiltration of E1 constructs 39 (type 1b) and 62 (type 3a). The proteins were labelled with NEM-biotin. Lane 1: start gelfiltration construct 39, lane 2: fraction 26 construct 39, lane 3: fraction 27 construct 39, lane 4: fraction 28 construct 39, lane 5: fraction 29 construct 39, lane 6: fraction 30 construct 39, lane 7: fraction 31 construct 39, lane 8: molecular weight marker, lane 9: start gelfiltration construct 62, lane 10: fraction 26 construct 62, lane 11: fraction 27 construct 62, lane 12: fraction 28 construct 62, lane 13: fraction 29 construct 62, lane 14: fraction 30 construct 62, lane 15: fraction 31 construct 62.

Figure 28: Silver staining of an SDS-PAGE gel of the gelfiltration fractions of vvHCV-39 (E1s, type 1b) and vvHCV-62 (E1s, type 3a) run under identical conditions as Figure 26. Lane 1: start gelfiltration construct 39, lane 2: fraction 26 construct 39, lane 3: fraction 27 construct 39, lane 4: fraction 28 construct 39, lane 5: fraction 29 construct 39, lane 6: fraction 30 construct 39, lane 7: fraction 31 construct 39, lane 8: molecular weight marker, lane 9: start gelfiltration construct 62, lane 10: fraction 26 construct 62, lane 11: fraction

27 construct 62, lane 12: fraction 28 construct 62, lane 13: fraction 29  
construct 62, lane 14: fraction 30 construct 62, lane 15: fraction 31 construct  
62.

5 Figure 29: Western Blot analysis with anti-E1 mouse monoclonal antibody 5E1A10  
giving a complete overview of the purification procedure. Lane 1: crude  
lysate, Lane 2: flow through of lentil chromatography, Lane 3: wash with  
Empigen BB after lentil chromatography, Lane 4: Eluate of lentil  
chromatography, Lane 5: Flow through during concentration of the lentil  
eluate, Lane 6: Pool of E1 after Size Exclusion Chromatography  
10 (gelfiltration).

Figure 30: OD<sub>280</sub> profile (continuous line) of the lentil lectin chromatography of E2  
protein from RK13 cells infected with vvHCV44. The dotted line represents  
the E2 reactivity as detected by ELISA (as in example 6).

15 Figure 31A: OD<sub>280</sub> profile (continuous line) of the lentil-lectin gelfiltration  
chromatography E2 protein pool from RK13 cells infected with vvHCV44 in  
which the E2 pool is applied immediately on the gelfiltration column (non-  
reduced conditions). The dotted line represents the E2 reactivity as detected  
by ELISA (as in example 6).

20 Figure 31B: OD<sub>280</sub> profile (continuous line) of the lentil-lectin gelfiltration  
chromatography E2 protein pool from RK13 cells infected with vvHCV44 in  
which the E2 pool was reduced and blocked according to Example 5.3  
(reduced conditions). The dotted line represents the E2 reactivity as detected  
by ELISA (as in example 6).

25 Figure 32: Ni<sup>2+</sup>-IMAC chromatography and ELISA reactivity of the E2 protein as  
expressed from vvHCV44 after gelfiltration under reducing conditions as  
shown in Figure 31B.

Figure 33: Silver staining of an SDS-PAGE of 0.5 µg of purified E2 protein recovered  
by a 200 mM imidazole elution step (lane 2) and a 30mM imidazole wash  
(lane 1) of the Ni<sup>2+</sup>-IMAC chromatography as shown in Figure 32.

30 Figure 34: OD profiles of a desalting step of the purified E2 protein recovered by 200  
mM imidazole as shown in Figure 33, intended to remove imidazole.

Figure 35A: Antibody levels to the different HCV antigens (Core 1, Core 2, E2HCVR,  
NS3) for NR and LTR followed during treatment and over a period of 6 to 12

months after treatment determined by means of the LIAscan method. The average values are indicated by the curves with the open squares.

Figure 35B: Antibody levels to the different HCV antigens (NS4, NS5, E1 and E2) for NR and LTR followed during treatment and over a period of 6 to 12 months after treatment determined by means of the LIAscan method. The average values are indicated by the curve with the open squares.

Figure 36: Average E1 antibody (E1Ab) and E2 antibody (E2Ab) levels in the LTR and NR groups.

Figure 37: Averages E1 antibody (E1Ab) levels for non-responders (NR) and long term responders (LTR) for type 1b and type 3a.

Figure 38: Relative map positions of the anti-E2 monoclonal antibodies.

Figure 39: Partial deglycosylation of HCV E1 envelope protein. The lysate of vvHCV10A-infected RK13 cells were incubated with different concentrations of glycosidases according to the manufacturer's instructions. Right panel: Glycopeptidase F (PNGase F). Left panel: Endoglycosidase H (Endo H).

Figure 40: Partial deglycosylation of HCV E2 envelope proteins. The lysate of vvHCV64-infected (E2) and vvHCV41-infected (E2s) RK13 cells were incubated with different concentrations of Glycopeptidase F (PNGase F) according to the manufacturer's instructions.

Figure 41: In vitro mutagenesis of HCV E1 glycoproteins. Map of the mutated sequences and the creation of new restriction sites.

Figure 42A: In vitro mutagenesis of HCV E1 glycoprotein (part 1). First step of PCR amplification.

Figure 42B: In vitro mutagenesis of HCV E1 glycoprotein (part 2). Overlap extension and nested PCR.

Figure 43: In vitro mutagenesis of HCV E1 glycoproteins. Map of the PCR mutated fragments (GLY-# and OVR-#) synthesized during the first step of amplification.

Figure 44A: Analysis of E1 glycoprotein mutants by Western blot expressed in HeLa (left) and RK13 (right) cells. Lane 1: wild type VV (vaccinia virus), Lane 2: original E1 protein (vvHCV-10A), Lane 3: E1 mutant Gly-1 (vvHCV-81),

Lane 4: E1 mutant Gly-2 (vvHCV-82), Lane 5: E1 mutant Gly-3 (vvHCV-83), Lane 6: E1 mutant Gly-4 (vvHCV-84), Lane 7: E1 mutant Gly-5 (vvHCV-85), Lane 8: E1 mutant Gly-6 (vvHCV-86).

5 Figure 44B: Analysis of E1 glycosylation mutant vaccinia viruses by PCR amplification/restriction. Lane 1: E1 (vvHCV-10A), *BspE* I, Lane 2: E1.GLY-1 (vvHCV-81), *BspE* I, Lane 4: E1 (vvHCV-10A), *Sac* I, Lane 5: E1.GLY-2 (vvHCV-82), *Sac* I, Lane 7: E1 (vvHCV-10A), *Sac* I, Lane 8: E1.GLY-3 (vvHCV-83), *Sac* I, Lane 10: E1 (vvHCV-10A), *Stu* I, Lane 11: E1.GLY-4 (vvHCV-84), *Stu* I, Lane 13: E1 (vvHCV-10A), *Sma* I, Lane 14: E1.GLY-5 (vvHCV-85), *Sma* I, Lane 16: E1 (vvHCV-10A), *Stu* I, Lane 17: E1.GLY-6 (vvHCV-86), *Stu* I, Lane 3 - 6 - 9 - 12 - 15 : Low Molecular Weight Marker, pBluescript SK+, *Msp* I.

10 Figure 45: SDS polyacrylamide gel electrophoresis of recombinant E2 expressed in *S. cerevisiae*. Inoculates were grown in leucine selective medium for 72 hrs. and diluted 1/15 in complete medium. After 10 days of culture at 28°C, medium samples were taken. The equivalent of 200 µl of culture supernatant concentrated by speedvac was loaded on the gel. Two independent transformants were analysed.

15 Figure 46: SDS polyacrylamide gel electrophoresis of recombinant E2 expressed in a glycosylation deficient *S. cerevisiae* mutant. Inoculae were grown in leucine selective medium for 72 hrs. and diluted 1/15 in complete medium. After 10 days of culture at 28°C, medium samples were taken. The equivalent of 350 µl of culture supernatant, concentrated by ion exchange chromatography, was loaded on the gel.

20 Figure 47: Profile of chimpanzees and immunization schedule.

Figure 48: Cellular response after 3 immunizations.

Figure 49: Evolution of cellular response upon repeated E1 immunizations.

Figure 50: Cellular response upon NS3 immunizations.

25 Figure 51: Stimulation index through week 28. The stimulation index (SI; cellular immune response) was obtained by culturing PBMC ( $10^5$  cells), drawn from the individuals before immunization (week 0), two weeks after the third immunization (week 8), before the booster immunization (week 26) and two weeks after the booster immunization (week 28), in the presence or absence



of 3  $\mu$ g of recombinant E1s or 2  $\mu$ g tetanos toxoid and determining the amount of tritiated thymidine incorporated in these cells during a pulse of 18 hours after 5 days of culture. The stimulation index is the ratio of thymidine incorporated in the cells cultured with envelope antigen versus the ones cultured without antigen. Samples of week 0 and 8 were determined in a first assay (A), while the samples of week 26 and 28 were determined in a second assay (B) in which the samples of week 0 were reanalyzed. Results are expressed as the geometric mean stimulation index of all 20 (A, experiment) or 19 (B, experiment) volunteers.

10 Figure 52:

Cytokine production of PBMCs. PBMC ( $10^5$  cells), drawn from the individuals before the booster immunization (week 26) and two weeks after the booster immunization (week 28), were cultured in the presence of 3  $\mu$ g of recombinant E1s (E1) or 2  $\mu$ g of tetanos toxoid (TT) or no antigen (BI). Cytokines were measured in the supernatant taken after 24 hours (interleukin-5) or after 120 hours (interferon-gamma) by means of ELISA. The stimulation index is the ratio of cytokine measured in the supernatants of cells cultured with envelope antigen versus the ones cultured without antigen. Results are expressed as the geometric mean of pg cytokine/ml secreted of all 19 volunteers. Samples with a cytokine amount below detection limit were assigned the value of the detection limit. Similarly samples with extremely high concentrations of cytokine out of the linear range of the assay were assigned the value of the limit of the linear range of the assay.

20 Figure 53:

Thymidine incorporation results. The stimulation index (cellular immune response) was obtained by culturing PBMC ( $3 \times 10^5$  cells), in the presence or absence of peptides and determining the amount of tritiated thymidine incorporated in these cells during a pulse after 5-6 days of culture. The stimulation index is the ratio of thymidine incorporated in the cells cultured with peptide versus the ones cultured without peptide. Results are expressed as individual values for vaccinated persons (top panel) or non vaccinated or controls (lower panel).

30 Figure 54:

Three-dimensional graph showing for the individual patients (each represented by a dot) the % change in ALT-levels (absolute change from

baseline) on the X-axis, the change in serum anti-E1 antibody levels (in mU/mL) on the Y-axis, and the change in Ishak fibrosis score on the Z-axis.

Figure 55: Three-dimensional graph showing for the individual patients (each represented by a dot) the % change in ALT-levels (absolute change from baseline) on the X-axis, the change in serum anti-E1 antibody levels (in mU/mL) on the Y-axis, and the change in Metavir fibrosis score on the Z-axis.

Figure 56: For each individual patient (represented by a dot or a "★"), the Ishak fibrosis score (on the X-axis) and the ALT values (on the Y-axis) are given. In panel A (top), the baseline (pre-vaccination) situation is given whereas panel B (bottom) is illustrating the situation at the time of taking of liver biopsies. The seven patients with the highest increase in serum anti-E1 antibody level are represented by a "★".

Figure 57: The influence of the age (on the X-axis) of each individual patient (represented by a dot or a "★") on the Ishak fibrosis score (on the Y-axis) is given. In panel A (top), the baseline (pre-vaccination) situation is given whereas panel B (bottom) is illustrating the situation at the time of taking of liver biopsies. The seven patients with the highest increase in serum anti-E1 antibody level are represented by a "★".

Table 1: Features of the respective clones and primers used for amplification for constructing the different forms of the E1 protein as despected in Example 1.

Table 2: Summary of Anti-E1 tests

Table 3: Synthetic peptides for competition studies

Table 4: Changes of envelope antibody levels over time.

Table 5: Difference between LTR and NR

Table 6: Competition experiments between murine E2 monoclonal antibodies

Table 7: Primers for construction of E1 glycosylation mutants

Table 8: Analysis of E1 glycosylation mutants by ELISA

Table 9: Profile of adjuvanted E1 Balb/c mice.

Table 10: Humoral responses: No. of immunizations required for different E1-antibodies levels.

Table 11: Chimpanzee antibody titers.

Table 12: Human antibody titers.

Table 13: Human antibody titers (8-28 weeks).

Table 14: Stimulation index (SI) of cultured PBMC, drawn from the individuals four weeks (W16) after the fourth immunization and two weeks (W26) after the fifth immunization in the presence or absence of 3  $\mu$ g of E1s. A stimulation index of >3 is considered a positive signal.

Table 15: Ishak grading of necro-inflammatory intensities for periportal hepatitis, confluent necrosis, focal inflammation, portal inflammation and the overall total inflammation grading. Scores are indicated as the change from baseline (mean and 95 % confidence interval) and the mean baseline- to end-values.

Table 16: Overview of frequencies (given as number of patients) of changes of a given baseline Metavir score (given in top row; Baseline 0 to 4) to a given Metavir score at the end of the second course of therapeutic E1s vaccination (given in left row; EOT 0 to 4). For instance, the "5" marked with a "\*" (i.e., "5\*") means that 5 patients had a baseline Metavir score of 1 and a Metavir score of 0 at the end of the second course treatment (EOT = end of treatment).

Table 17: Correlation between serum anti-E1 antibody levels induced by therapeutic E1s vaccination and change in overall Ishak scores. Given are the number of patients corresponding to the possible criteria as outlined in the Table.

The pgptATA18 vaccinia recombination plasmid is a modified version of pATA18 (Stunnenberg et al, 1988) with an additional insertion containing the *E. coli* xanthine guanine phosphoribosyl transferase gene under the control of the vaccinia virus I3 intermediate promoter (Figure 1). The plasmid pgsATA18 was constructed by inserting an oligonucleotide linker with SEQ ID NO 1/94, containing stop codons in the three reading frames, into the *Pst* I and *Hind*III-cut pATA18 vector. This created an extra *Pac* I restriction site (Figure 2). The original *Hind*III site was not restored.

5' G GCATGC AAGCTT AATTAATT 3' (SEQ ID NO:1)  
3' ACGTC CGTACG TTCGAA TTAATTAA TCGA 5' (SEQ ID NO:94)

In order to facilitate rapid and efficient purification by means of  $\text{Ni}^{2+}$  chelation of engineered histidine stretches fused to the recombinant proteins, the vaccinia recombination vector pMS66 was designed to express secreted proteins with an additional carboxy-terminal histidine tag. An oligonucleotide linker with SEQ ID NO 2/95, containing unique sites for 3 restriction enzymes generating blunt ends (*Sma* I, *Stu* I and *Pml* I/*Bbr* PI) was synthesized in such a way that the carboxy-terminal end of any cDNA could be inserted in frame with a sequence encoding the protease factor Xa cleavage site followed by a nucleotide sequence encoding 6 histidines and 2 stop codons (a new *Pac* I restriction site was also created downstream the 3' end). This oligonucleotide with SEQ ID NO 2/95 was introduced between the *Xma* I and *Pst* I sites of pgpATA18 (Figure 3).

5'-CCGGG GAGGCCTGCACGTGATCGAGGGCAGACACCATCACCACCA/  
3'-C CTCCGGACGTGCACTAGCTCCCGTCTGTGGTAGTGGTGGT/  
*Xma*I

NO:2) /TCACTAATAGTTAATTAA CTGCA-3' (SEQ ID

NO:95) /AGTGATTATCAATTAATT G-5' (SEQ ID

35

Plasmid pgptATA-18 contained within *Escherichia coli* MC1061 (lambda) has been deposited under the terms of the Budapest Treaty at BCCM/LMBP (Belgian Coordinated Collections of microorganisms/Laboratorium voor Moleculaire Biologie - Plasmidencollectie, Universiteit Gent, K.L. Ledeganckstraat 35, B-9000 Ghent, Belgium), and bears accession number LMBP4486. Said deposit was made on January 9, 2002.

## **Example 2. Construction of HCV recombinant plasmids**

### **2.1. Constructs encoding different forms of the E1 protein**

10 Polymerase Chain Reaction (PCR) products were derived from the serum samples by RNA preparation and subsequent reverse-transcription and PCR as described previously (Stuyver et al., 1993b). Table 1 shows the features of the respective clones and the primers used for amplification. The PCR fragments were cloned into the Sma I-cut pSP72 (Promega) plasmids. The following clones were selected for insertion into vaccinia recombination  
15 vectors: HCCI9A (SEQ ID NO:3), HCCI10A (SEQ ID NO:5), HCCI11A (SEQ ID NO:7), HCCI12A (SEQ ID NO:9), HCCI13A (SEQ ID NO:11), and HCCI17A (SEQ ID NO:13) as depicted in Figure 21. cDNA fragments containing the E1-coding regions were cleaved by EcoRI and HindIII restriction from the respective pSP72 plasmids and inserted into the  
20 downstream of the 11K vaccinia virus late promoter. The respective plasmids were designated pvHCV-9A, pvHCV-10A, pvHCV-11A, pvHCV-12A, pvHCV-13A and pvHCV-17A, of which pvHCV-11A is shown in Figure 4.

### **2.2. Hydrophobic region E1 deletion mutants**

25 Clone HCCI37, containing a deletion of codons Asp264 to Val287 (nucleotides 790 to 861, region encoding hydrophobic domain I) was generated as follows: 2 PCR fragments were generated from clone HCCI10A with primer sets HCPPr52 (SEQ ID NO:16)/HCPPr107 (SEQ ID NO:19) and HCPPr108 (SEQ ID NO:20)/HCPPr54 (SEQ ID NO:18). These primers are shown in Figure 21. The two PCR fragments were purified from agarose gel after  
30 electrophoresis and 1 ng of each fragment was used together as template for PCR by means of primers HCPPr52 (SEQ ID NO:16) and HCPPr54 (SEQ ID NO:18). The resulting fragment was cloned into the Sma I-cut pSP72 vector and clones containing the deletion were readily identified because of the deletion of 24 codons (72 base pairs). Plasmid pSP72HCCI37

containing clone HCC137 (SEQ ID 15) was selected. A recombinant vaccinia plasmid containing the full-length E1 cDNA lacking hydrophobic domain I was constructed by inserting the HCV sequence surrounding the deletion (fragment cleaved by Xma I and BamH I from the vector pSP72-HCC137) into the Xma I-Bam H I sites of the vaccinia plasmid pvHCV-10A. The resulting plasmid was named pvHCV-37. After confirmatory sequencing, the amino-terminal region containing the internal deletion was isolated from this vector pvHCV-37 (cleavage by EcoR I and BstE II) and reinserted into the Eco RI and Bst EII-cut pvHCV-11A plasmid. This construct was expected to express an E1 protein with both hydrophobic domains deleted and was named pvHCV-38. The E1-coding region of clone HCC138 is represented by SEQ ID NO:23.

As the hydrophilic region at the E1 carboxyterminus (theoretically extending to around amino acids 337-340) was not completely included in construct pvHCV-38, a larger E1 region lacking hydrophobic domain I was isolated from the pvHCV-37 plasmid by EcoR I/Bam HI cleavage and cloned into an EcoRI/BamHI-cut pgsATA-18 vector. The resulting plasmid was named pvHCV-39 and contained clone HCC139 (SEQ ID NO:25). The same fragment was cleaved from the pvHCV-37 vector by BamH I (of which the sticky ends were filled with Klenow DNA Polymerase I (Boehringer)) and subsequently by EcoR I (5' cohesive end). This sequence was inserted into the EcoRI and Bbr PI-cut vector pMS-66. This resulted in clone HCC140 (SEQ ID NO:27) in plasmid pvHCV-40, containing a 6 histidine tail at its carboxy-terminal end.

### 2.3. E1 of other genotypes

Clone HCC162 (SEQ ID NO:29) was derived from a type 3a-infected patient with chronic hepatitis C (serum BR36, clone BR36-9-13, SEQ ID NO:19 in WO 94/25601, and see also Stuyver et al. 1993a) and HCC163 (SEQ ID NO:31) was derived from a type 5a-infected child with post-transfusion hepatitis (serum BE95, clone PC-4-1, SEQ ID NO:45 in WO 94/25601).

### 2.4. E2 constructs

The HCV E2 PCR fragment 22 was obtained from serum BE11 (genotype 1b) by means of primers HCP109 (SEQ ID NO:33) and HCP172 (SEQ ID NO:34) using techniques of RNA preparation, reverse-transcription and PCR, as described in Stuyver et al., 1993b, and the fragment was cloned into the Sma I-cut pSP72 vector. Clone HCC122A (SEQ ID

NO:35) was cut with NcoI/AlwNI or by BamHI/AlwNI and the sticky ends of the fragments were blunted (NcoI and BamHI sites with Klenow DNA Polymerase I (Boehringer), and AlwNI with T4 DNA polymerase (Boehringer)). The BamHI/AlwNI cDNA fragment was then inserted into the vaccinia pgsATA-18 vector that had been linearized by EcoR I and Hind III cleavage and of which the cohesive ends had been filled with Klenow DNA Polymerase (Boehringer). The resulting plasmid was named pvHCV-41 and encoded the E2 region from amino acids Met347 to Gln673, including 37 amino acids (from Met347 to Gly383) of the E1 protein that can serve as signal sequence. The same HCV cDNA was inserted into the EcoR I and Bbr PI-cut vector pMS66, that had subsequently been blunt ended with Klenow DNA Polymerase. The resulting plasmid was named pvHCV-42 and also encoded amino acids 347 to 683. The NcoI/AlwNI fragment was inserted in a similar way into the same sites of pgsATA-18 (pvHCV-43) or pMS-66 vaccinia vectors (pvHCV-44). pvHCV-43 and pvHCV-44 encoded amino acids 364 to 673 of the HCV polyprotein, of which amino acids 364 to 383 were derived from the natural carboxyterminal region of the E1 protein encoding the signal sequence for E2, and amino acids 384 to 673 of the mature E2 protein.

#### 2.5. Generation of recombinant HCV-vaccinia viruses

Rabbit kidney RK13 cells (ATCC CCL 37), human osteosarcoma 143B thymidine kinase deficient (TK<sup>-</sup>) (ATCC CRL 8303), HeLa (ATCC CCL 2), and Hep G2 (ATCC HB 8065) cell lines were obtained from the American Type Culture Collection (ATCC, Rockville, Md, USA). The cells were grown in Dulbecco's modified Eagle medium (DMEM) supplemented with 10 % foetal calf serum, and with Earle's salts (EMEM) for RK13 and 143 B (TK<sup>-</sup>), and with glucose (4 g/l) for Hep G2. The vaccinia virus WR strain (Western Reserve, ATTC VR119) was routinely propagated in either 143B or RK13 cells, as described previously (Panicali & Paoletti, 1982; Piccini et al., 1987; Mackett et al., 1982, 1984, and 1986). A confluent monolayer of 143B cells was infected with wild type vaccinia virus at a multiplicity of infection (m.o.i.) of 0.1 (= 0.1 plaque forming unit (PFU) per cell). Two hours later, the vaccinia recombination plasmid was transfected into the infected cells in the form of a calcium phosphate coprecipitate containing 500 ng of the plasmid DNA to allow homologous recombination (Graham & van der Eb, 1973; Mackett et al., 1985). Recombinant viruses expressing the *E.coli* xanthine-guanine phosphoribosyl transferase (gpt) protein were selected on rabbit kidney RK13 cells incubated in selection medium (EMEM

containing 25 µg/ml mycophenolic acid (MPA), 250 µg/ml xanthine, and 15 µg/ml hypoxanthine; Falkner and Moss, 1988; Janknecht et al, 1991). Single recombinant viruses were purified on fresh monolayers of RK13 cells under a 0.9% agarose overlay in selection medium. Thymidine kinase deficient (TK<sup>-</sup>) recombinant viruses were selected and then plaque purified on fresh monolayers of human 143B cells (TK<sup>-</sup>) in the presence of 25 µg/ml 5-bromo-2'-deoxyuridine. Stocks of purified recombinant HCV-vaccinia viruses were prepared by infecting either human 143 B or rabbit RK13 cells at an m.o.i. of 0.05 (Mackett et al, 1988). The insertion of the HCV cDNA fragment in the recombinant vaccinia viruses was confirmed on an aliquot (50 µl) of the cell lysate after the MPA selection by means of PCR with the primers used to clone the respective HCV fragments (see Table 1). The recombinant vaccinia-HCV viruses were named according to the vaccinia recombination plasmid number, e.g. the recombinant vaccinia virus vvHCV-10A was derived from recombining the wild type WR strain with the pvHCV-10A plasmid.

#### 15 **Example 3: infection of cells with recombinant vaccinia viruses**

A confluent monolayer of RK13 cells was infected at a m.o.i. of 3 with the recombinant HCV-vaccinia viruses as described in example 2. For infection, the cell monolayer was washed twice with phosphate-buffered saline pH 7.4 (PBS) and the recombinant vaccinia virus stock was diluted in MEM medium. Two hundred µl of the virus solution was added per 10<sup>6</sup> cells such that the m.o.i. was 3, and incubated for 45 min at 24°C. The virus solution was aspirated and 2 ml of complete growth medium (see example 2) was added per 10<sup>6</sup> cells. The cells were incubated for 24 hr at 37°C during which expression of the HCV proteins took place.

#### 25 **Example 4: Analysis of recombinant proteins by means of western blotting**

The infected cells were washed two times with PBS, directly lysed with lysis buffer (50 mM Tris.HCl pH 7.5, 150 mM NaCl, 1% Triton X-100, 5 mM MgCl<sub>2</sub>, 1 µg/ml aprotinin (Sigma, Bornem, Belgium)) or detached from the flasks by incubation in 50 mM Tris.HCL pH 7.5/ 10 mM EDTA/ 150 mM NaCl for 5 min, and collected by centrifugation (5 min at 1000g). The cell pellet was then resuspended in 200 µl lysis buffer (50 mM Tris.HCL pH 8.0, 2 mM EDTA, 150 mM NaCl, 5 mM MgCl<sub>2</sub> aprotinin, 1% Triton X-100) per 10<sup>6</sup> cells.

The cell lysates were cleared for 5 min at 14,000 rpm in an Eppendorf centrifuge to remove the insoluble debris. Proteins of 20 µl lysate were separated by means of sodium



dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE). The proteins were then electro-transferred from the gel to a nitrocellulose sheet (Amersham) using a Hoefer HSI transfer unit cooled to 4°C for 2 hr at 100 V constant voltage, in transfer buffer (25 mM Tris.HCl pH 8.0, 192 mM glycine, 20% (v/v) methanol). Nitrocellulose filters were blocked with Blotto (5 % (w/v) fat-free instant milk powder in PBS; Johnson et al., 1981) and incubated with primary antibodies diluted in Blotto/0.1 % Tween 20. Usually, a human negative control serum or serum of a patient infected with HCV were 200 times diluted and preincubated for 1 hour at room temperature with 200 times diluted wild type vaccinia virus-infected cell lysate in order to decrease the non-specific binding. After washing with Blotto/0.1% Tween 20, the nitrocellulose filters were incubated with alkaline phosphatase substrate solution diluted in Blotto/0.1 % Tween 20. After washing with 0.1% Tween 20 in PBS, the filters were incubated with alkaline phosphatase substrate solution (100 mM Tris.HCl pH 9.5, 100 mM NaCl, 5 mM MgCl<sub>2</sub>, 0,38 µg/ml nitroblue tetrazolium, 0.165 µg/ml 5-bromo-4-chloro-3-indolylphosphate). All steps, except the electrotransfer, were performed at room temperature.

### **Example 5: Purification of recombinant E1 or E2 protein**

#### **5.1. Lysis**

Infected RK13 cells (carrying E1 or E2 constructs) were washed 2 times with phosphate-buffered saline (PBS) and detached from the culture recipients by incubation in PBS containing 10 mM EDTA. The detached cells were washed twice with PBS and 1 ml of lysis buffer (50 mM Tris.HCl pH 7.5, 150 mM NaCl, 1% Triton X-100, 5 mM MgCl<sub>2</sub>, 1 µg/ml aprotinin (Sigma, Bornem, Belgium) containing 2 mM biotinylated N-ethylmaleimide (biotin-NEM) (Sigma) was added per 10<sup>5</sup> cells at 4°C. This lysate was homogenized with a type B douncer and left at room temperature for 0.5 hours. Another 5 volumes of lysis buffer containing 10 mM N-ethylmaleimide (NEM, Aldrich, Bornem, Belgium) was added to the primary lysate and the mixture was left at room temperature for 15 min. Insoluble cell debris was cleared from the solution by centrifugation in a Beckman JA-14 rotor at 14,000 rpm (30100 g at  $r_{max}$ ) for 1 hour at 4°C.

#### **5.2. Lectin Chromatography**

The cleared cell lysate was loaded at a rate of 1ml/min on a 0.8 by 10 cm Lentil-lectin

Sepharose 4B column (Pharmacia) that had been equilibrated with 5 column volumes of lysis buffer at a rate of 1ml/min. The lentil-lectin column was washed with 5 to 10 column volumes of buffer 1 (0.1M potassium phosphate pH 7.3, 500 mM KCl, 5% glycerol, 1 mM 6-NH<sub>2</sub>-hexanoic acid, 1 mM MgCl<sub>2</sub>, and 1% DecylPEG (KWANT, Bedum, The Netherlands). In some experiments, the column was subsequently washed with 10 column volumes of buffer 1 containing 0.5% Empigen-BB (Calbiochem, San Diego, CA, USA) instead of 1% DecylPEG. The bound material was eluted by applying elution buffer (10 mM potassium phosphate pH 7.3, 5% glycerol, 1 mM hexanoic acid, 1mM MgCl<sub>2</sub>, 0.5% Empigen-BB, and 0.5 M  $\alpha$ -methyl-mannopyranoside). The eluted material was fractionated and fractions were screened for the presence of E1 or E2 protein by means of ELISA as described in example 6. Figure 22 shows ELISA results obtained from lentil lectin eluate fractions of 4 different E1 purifications of cell lysates infected with vvHCV39 (type 1b), vvHCV40 (type 1b), vvHCV62 (type 3a), and vvHCV63 (type 5a). Figure 23 shows the profiles obtained from the values shown in Figure 22. These results show that the lectin affinity column can be employed for envelope proteins of the different types of HCV.

### 5.3. Concentration and partial reduction

The E1- or E2-positive fractions were pooled and concentrated on a Centricon 30 kDa (Amicon) by centrifugation for 3 hours at 5,000 rpm in a Beckman JA-20 rotor at 4°C. In some experiments the E1- or E2-positive fractions were pooled and concentrated by nitrogen evaporation. An equivalent of  $3.10^8$  cells was concentrated to approximately 200  $\mu$ l. For partial reduction, 30% Empigen-BB (Calbiochem, San Diego, CA, USA) was added to this 200  $\mu$ l to a final concentration of 3.5 %, and 1M DTT in H<sub>2</sub>O was subsequently added to a final concentration of 1.5 to 7.5 mM and incubated for 30 min at 37 °C. NEM (1M in dimethylsulphoxide) was subsequently added to a final concentration of 50 mM and left to react for another 30 min at 37°C to block the free sulphydryl groups.

### 5.4. Gel filtration chromatography

A Superdex-200 HR 10/20 column (Pharmacia) was equilibrated with 3 column volumes PBS/3% Empigen-BB. The reduced mixture was injected in a 500  $\mu$ l sample loop of the Smart System (Pharmacia) and PBS/3% Empigen-BB buffer was added for gelfiltration. Fractions of 250  $\mu$ l were collected from V<sub>0</sub> to V<sub>i</sub>. The fractions were screened for the

presence of E1 or E2 protein as described in example 6.

Figure 24 shows ELISA results obtained from fractions obtained after gelfiltration chromatography of 4 different E1 purifications of cell lysates infected with vvHCV39 (type 1b), vvHCV40 (type 1b), vvHCV62 (type 3a), and vvHCV63 (type 5a). Figure 25 shows the profiles obtained from purifications of E1 proteins of types 1b, 3a, and 5a (from RK13 cells infected with vvHCV39, vvHCV62, and vvHCV63, respectively; purified on lentil lectin and reduced as in the previous examples). The peaks indicated with '1', '2', and '3', represent pure E1 protein peaks (E1 reactivity mainly in fractions 26 to 30). These peaks show very similar molecular weights of approximately 70 kDa, corresponding to dimeric E1 protein. Other peaks in the three profiles represent vaccinia virus and/or cellular proteins which could be separated from E1 only because of the reduction step as outlined in example 5.3. and because of the subsequent gelfiltration step in the presence of the proper detergent. As shown in Figure 26 pool 1 (representing fractions 10 to 17) and pool 2 (representing fractions 18 to 25) contain contaminating proteins not present in the E1 pool (fractions 26 to 30). The E1 peak fractions were ran on SDS/PAGE and blotted as described in example 4. Proteins labelled with NEM-biotin were detected by streptavidin-alkaline phosphatase as shown in Figure 27. It can be readily observed that, amongst others, the 29 kDa and 45kDa contaminating proteins present before the gelfiltration chromatography (lane 1) are only present at very low levels in the fractions 26 to 30. The band at approximately 65kDa represents the E1 dimeric form that could not be entirely disrupted into the monomeric E1 form. Similar results were obtained for the type 3a E1 protein (lanes 10 to 15), which shows a faster mobility on SDS/PAGE because of the presence of only 5 carbohydrates instead of 6. Figure 28 shows a silver stain of an SDS/PAGE gel run in identical conditions as in Figure 26. A complete overview of the purification procedure is given in Figure 29.

The presence of purified E1 protein was further confirmed by means of western blotting as described in example 4. The dimeric E1 protein appeared to be non-aggregated and free of contaminants. The subtype 1b E1 protein purified from vvHCV40-infected cells according to the above scheme was aminoterminally sequenced on an 477 Perkin-Elmer sequencer and appeared to contain a tyrosine as first residue. This confirmed that the E1 protein had been cleaved by the signal peptidase at the correct position (between A191 and Y192) from its signal sequence. This confirms the finding of Hijikata et al. (1991) that the aminotermius of the mature E1 protein starts at amino acid position 192.

### 5.5. Purification of the E2 protein

The E2 protein (amino acids 384 to 673) was purified from RK13 cells infected with vvHCV44 as indicated in Examples 5.1 to 5.4. Figure 30 shows the OD<sub>280</sub> profile (continuous line) of the lentil lectin chromatography. The dotted line represents the E2 reactivity as detected by ELISA (see example 6). Figure 31 shows the same profiles obtained from gelfiltration chromatography of the lentil-lectin E2 pool (see Figure 30), part of which was reduced and blocked according to the methods as set out in example 5.3., and part of which was immediately applied to the column. Both parts of the E2 pool were run on separate gelfiltration columns. It could be demonstrated that E2 forms covalently-linked aggregates with contaminating proteins if no reduction has been performed. After reduction and blocking, the majority of contaminating proteins segregated into the V<sub>0</sub> fraction. Other contaminating proteins copurified with the E2 protein, were not covalently linked to the E2 protein any more because these contaminants could be removed in a subsequent step. Figure 32 shows an additional Ni<sup>2+</sup>-IMAC purification step carried out for the E2 protein purification. This affinity purification step employs the 6 histidine residues added to the E2 protein as expressed from vvHCV44. Contaminating proteins either run through the column or can be removed by a 30 mM imidazole wash. Figure 33 shows a silver-stained SDS/PAGE of 0.5 µg of purified E2 protein and a 30 mM imidazole wash. The pure E2 protein could be easily recovered by a 200 mM imidazole elution step. Figure 34 shows an additional desalting step intended to remove imidazole and to be able to switch to the desired buffer, e.g. PBS, carbonate buffer, saline.

Starting from about 50,000 cm<sup>2</sup> of RK13 cells infected with vvHCV11A (or vvHCV40) for the production of E1 or vvHCV41, vvHCV42, vvHCV43, or vvHCV44 for production of E2 protein, the procedures described in examples 5.1 to 5.5 allow the purification of approximately 1.3 mg of E1 protein and 0.6 mg of E2 protein.

It should also be remarked that secreted E2 protein (constituting approximately 30-40%, 60-70% being in the intracellular form) is characterized by aggregate formation (contrary to expectations). The same problem is thus posed to purify secreted E2. The secreted E2 can be purified as disclosed above.

### Example 6: ELISA for the detection of anti-E1 or anti-E2 antibodies or for the detection of E1 or E2 proteins

Maxisorb microwell plates (Nunc, Roskilde, Denmark) were coated with 1 volume

(e.g. 50  $\mu$ l or 100  $\mu$ l or 200  $\mu$ l) per well of a 5  $\mu$ g/ml solution of Streptavidin (Boehringer Mannheim) in PBS for 16 hours at 4°C or for 1 hour at 37°C. Alternatively, the wells were coated with 1 volume of 5  $\mu$ g/ml of *Galanthus nivalis* agglutinin (GNA) in 50 mM sodium carbonate buffer pH 9.6 for 16 hours at 4°C or for 1 hour at 37°C. In the case of coating with

5 GNA, the plates were washed 2 times with 400  $\mu$ l of Washing Solution of the Innatest HCV Ab III kit (Innogenetics, Belgium). Unbound coating surfaces were blocked with 1.5 to 2 volumes of blocking solution (0.1% casein and 0.1% NaN<sub>3</sub> in PBS) for 1 hour at 37°C or for 16 hours at 4°C. Blocking solution was aspirated. Purified E1 or E2 was diluted to 100-1000 ng/ml (concentration measured at A = 280 nm) or column fractions to be screened for E1 or

10 E2 (see example 5), or E1 or E2 in non-purified cell lysates (example 5.1.) were diluted 20 times in blocking solution, and 1 volume of the E1 or E2 solution was added to each well and incubated for 1 hour at 37°C on the Streptavidin- or GNA-coated plates. The microwells were washed 3 times with 1 volume of Washing Solution of the Innatest HCV Ab III kit (Innogenetics, Belgium). Serum samples were diluted 20 times or monoclonal anti-E1 or

15 anti-E2 antibodies were diluted to a concentration of 20 ng/ml in Sample Diluent of the Innatest HCV Ab III kit and 1 volume of the solution was left to react with the E1 or E2 protein for 1 hour at 37°C. The microwells were washed 5 times with 400  $\mu$ l of Washing Solution of the Innatest HCV Ab III kit (Innogenetics, Belgium). The bound antibodies were detected by incubating each well for 1 hour at 37°C with a goat anti-human or anti-mouse

20 IgG, peroxidase-conjugated secondary antibody (DAKO, Glostrup, Denmark) diluted 1/80,000 in 1 volume of Conjugate Diluent of the Innatest HCV Ab III kit (Innogenetics, Belgium), and color development was obtained by addition of substrate of the Innatest HCV Ab III kit (Innogenetics, Belgium) diluted 100 times in 1 volume of Substrate Solution of the Innatest HCV Ab III kit (Innogenetics, Belgium) for 30 min at 24°C after washing of the

25 plates 3 times with 400  $\mu$ l of Washing Solution of the Innatest HCV Ab III kit (Innogenetics, Belgium).

#### **Example 7: Follow up of patient groups with different clinical profiles**

##### **30 7.1. Monitoring of anti-E1 and anti-E2 antibodies**

The current hepatitis C virus (HCV) diagnostic assays have been developed for screening and confirmation of the presence of HCV antibodies. Such assays do not seem to provide information useful for monitoring of treatment or for prognosis of the outcome of

disease. However, as is the case for hepatitis B, detection and quantification of anti-envelope antibodies may prove more useful in a clinical setting. To investigate the possibility of the use of anti-E1 antibody titer and anti-E2 antibody titer as prognostic markers for outcome of hepatitis C disease, a series of IFN- $\alpha$  treated patients with long-term sustained response (defined as patients with normal transaminase levels and negative HCV-RNA test (PCR in the 5' non-coding region) in the blood for a period of at least 1 year after treatment) was compared with patients showing no response or showing biochemical response with relapse at the end of treatment.

A group of 8 IFN- $\alpha$  treated patients with long-term sustained response (LTR, follow up 1 to 3.5 years, 3 type 3a and 5 type 1b) was compared with 9 patients showing non-complete responses to treatment (NR, follow up 1 to 4 years, 6 type 1b and 3 type 3a). Type 1b (vvHCV-39, see example 2.5.) and 3a E1 (vvHCV-62, see example 2.5.) proteins were expressed by the vaccinia virus system (see examples 3 and 4) and purified to homogeneity (example 5). The samples derived from patients infected with a type 1b hepatitis C virus were tested for reactivity with purified type 1b E1 protein, while samples of a type 3a infection were tested for reactivity of anti-type 3a E1 antibodies in an ELISA as described in example 6. The genotypes of hepatitis C viruses infecting the different patients were determined by means of the Inno-LiPA genotyping assay (Innogenetics, Belgium). Figure 5 shows the anti-E1 signal-to-noise ratios of these patients followed during the course of interferon treatment and during the follow-up period after treatment. LTR cases consistently showed rapidly declining anti-E1 levels (with complete negativation in 3 cases), while anti-E1 levels of NR cases remained approximately constant. Some of the obtained anti-E1 data are shown in Table 2 as average S/N ratios  $\pm$  SD (mean anti-E1 titer). The anti-E1 titer could be deduced from the signal to noise ratio as show in Figures 5, 6, 7, and 8.

Already at the end of treatment, marked differences could be observed between the 2 groups. Anti-E1 antibody titers had decreased 6.9 times in LTR but only 1.5 times in NR. At the end of follow up, the anti-E1 titers had declined by a factor of 22.5 in the patients with sustained response and even slightly increased in NR. Therefore, based on these data, decrease of anti-E1 antibody levels during monitoring of IFN- $\alpha$  therapy correlates with long-term, sustained response to treatment. The anti-E1 assay may be very useful for prognosis of long-term response to IFN treatment, or to treatment of the hepatitis C disease in general.

This finding was not expected. On the contrary, the inventors had expected the anti-E1 antibody levels to increase during the course of IFN treatment in patients with long term

response. As is the case for hepatitis B, the virus is cleared as a consequence of the seroconversion for anti-HBsAg antibodies. Also in many other virus infections, the virus is eliminated when anti-envelope antibodies are raised. However, in the experiments of the present invention, anti-E1 antibodies clearly decreased in patients with a long-term response to treatment, while the antibody-level remained approximately at the same level in non-responding patients. Although the outcome of these experiments was not expected, this non-obvious finding may be very important and useful for clinical diagnosis of HCV infections. As shown in Figures 9, 10, 11, and 12, anti-E2 levels behaved very differently in the same patients studied and no obvious decline in titers was observed as for anti-E1 antibodies. Figure 35 gives a complete overview of the pilot study.

As can be deduced from Table 2, the anti-E1 titers were on average at least 2 times higher at the start of treatment in long term responders compared with incomplete responders to treatment. Therefore, measuring the titer of anti-E1 antibodies at the start of treatment, or monitoring the patient during the course of infection and measuring the anti-E1 titer, may become a useful marker for clinical diagnosis of hepatitis C. Furthermore, the use of more defined regions of the E1 or E2 proteins may become desirable, as shown in example 7.3.

#### 7.2. Analysis of E1 and E2 antibodies in a larger patient cohort

The pilot study lead the inventors to conclude that, in case infection was completely cleared, antibodies to the HCV envelope proteins changed more rapidly than antibodies to the more conventionally studied HCV antigens, with E1 antibodies changing most vigorously. We therefore included more type 1b and 3a-infected LTR and further supplemented the cohort with a matched series of NR, such that both groups included 14 patients each. Some partial responders (PR) and responders with relapse (RR) were also analyzed.

Figure 36 depicts average E1 antibody (E1Ab) and E2 antibody (E2Ab) levels in the LTR and NR groups and Tables 4 and 5 show the statistical analyses. In this larger cohort, higher E1 antibody levels before IFN- $\alpha$  therapy were associated with LTR ( $P < 0.03$ ). Since much higher E1 antibody levels were observed in type 3a-infected patients compared with type 1b-infected patients (Figure 37), the genotype was taken into account (Table 4). Within the type 1b-infected group, LTR also had higher E1 antibody levels than NR at the initiation of treatment [ $P < 0.05$ ]; the limited number of type 3a-infected NR did not allow statistical analysis.

Of antibody levels monitored in LTR during the 1.5-year follow up period, only E1

antibodies cleared rapidly compared with levels measured at initiation of treatment [ $P = 0.0058$ , end of therapy;  $P = 0.0047$  and  $P = 0.0051$  at 6 and 12 months after therapy, respectively]. This clearance remained significant within type 1- or type 3-infected LTR (average  $P$  values  $< 0.05$ ). These data confirmed the initial finding that E1Ab levels decrease rapidly in the early phase of resolution. This feature seems to be independent of viral genotype. In NR, PR, or RR, no changes in any of the antibodies measured were observed throughout the follow up period. In patients who responded favourably to treatment with normalization of ALT levels and HCV-RNA negative during treatment, there was a marked difference between sustained responders (LTR) and responders with a relapse (RR). In contrast to LTR, RR did not show any decreasing E1 antibody levels, indicating the presence of occult HCV infection that could neither be demonstrated by PCR or other classical techniques for detection of HCV-RNA, nor by raised ALT levels. The minute quantities of viral RNA, still present in the RR group during treatment, seemed to be capable of anti-E1 B cell stimulation. Anti-E1 monitoring may therefore not only be able to discriminate LTR from NR, but also from RR.

### 7.3. Monitoring of antibodies of defined regions of the E1 protein

Although the molecular biological approach of identifying HCV antigens resulted in unprecedented breakthrough in the development of viral diagnostics, the method of immune screening of  $\lambda$ gt11 libraries predominantly yielded linear epitopes dispersed throughout the core and non-structural regions, and analysis of the envelope regions had to await cloning and expression of the E1/E2 region in mammalian cells. This approach sharply contrasts with many other viral infections of which epitopes to the envelope regions had already been mapped long before the deciphering of the genomic structure. Such epitopes and corresponding antibodies often had neutralizing activity useful for vaccine development and/or allowed the development of diagnostic assays with clinical or prognostic significance (e.g. antibodies to hepatitis B surface antigen).

As no HCV vaccines or tests allowing clinical diagnosis and prognosis of hepatitis C disease are available today, the characterization of viral envelope regions exposed to immune surveillance may significantly contribute to new directions in HCV diagnosis and prophylaxis.

Several 20-mer peptides (Table 3) that overlapped each other by 8 amino acids, were synthesized according to a previously described method (EP-A-0 489 968) based on the HC-



J1 sequence (Okamoto et al., 1990). None of these, except peptide env35 (also referred to as E1-35), was able to detect antibodies in sera of approximately 200 HCV cases. Only 2 sera reacted slightly with the env35 peptide. However, by means of the anti-E1 ELISA as described in example 6, it was possible to discover additional epitopes as follows: The anti-  
 5 E1 ELISA as described in example 6 was modified by mixing 50 µg/ml of E1 peptide with the 1/20 diluted human serum in sample diluent. Figure 13 shows the results of reactivity of human sera to the recombinant E1 (expressed from vvHCV-40) protein, in the presence of single or of a mixture of E1 peptides. While only 2% of the sera could be detected by means of E1 peptides coated on strips in a Line Immunoassay format, over half of the sera contained  
 10 anti-E1 antibodies which could be competed by means of the same peptides, when tested on the recombinant E1 protein. Some of the murine monoclonal antibodies obtained from Balb/C mice after injection with purified E1 protein were subsequently competed for reactivity to E1 with the single peptides (Figure 14). Clearly, the region of env53 contained the predominant epitope, as the addition of env53 could substantially compete reactivity of  
 15 several sera with E1, and antibodies to the env31 region were also detected. This finding was surprising, since the env53 and env31 peptides had not shown any reactivity when coated directly to the solid phase.

Therefore peptides were synthesized using technology described by applicant previously (in WO 93/18054). The following peptides were synthesized:

20 peptide env35A-biotin

NH<sub>2</sub>-SNSSEAADMIMHTPGCV-GKbiotin (SEQ ID NO:51)

spanning amino acids 208 to 227 of the HCV polyprotein in the E1 region

peptide biotin-env53 (epitope A')

biotin-GG-ITGHRMAWDMMMNWSPTTAL-COOH (SEQ ID NO:52)

25 spanning amino acids to 313 of 332 of the HCV polyprotein in the E1 region

peptide 1bE1 ('epitope B')

H<sub>2</sub>N-YEVRNVSGIYHVTNDCSNSSIVYEAADMIMHTPGCGK -biotin  
 (SEQ ID NO:53)

spanning amino acids 192 to 228 of the HCV polyprotein in the E1 region

30 and compared with the reactivities of peptides E1a-BB (biotin-GG-TPTVATRDGKLPATQLRRHIDLL, SEQ ID NO:54) and E1b-BB (biotin-GG-TPTLAARDASVPTTTIRRHVDLL, SEQ ID NO:55) which are derived from the same region of sequences of genotype 1a and 1b respectively and which have been described at the

IXth international virology meeting in Glasgow, 1993 ('epitope C'). Reactivity of a panel of HCV sera was tested on epitopes A, B and C and epitope B was also compared with env35A (of 47 HCV-positive sera, 8 were positive on epitope B and none reacted with env35A). Reactivity towards epitopes A, B, and C was tested directly to the biotinylated peptides (50  
5 µg/ml) bound to streptavidin-coated plates as described in example 6. Clearly, epitopes A and B were most reactive while epitopes C and env35A-biotin were much less reactive. The same series of patients that had been monitored for their reactivity towards the complete E1 protein (example 7.1.) was tested for reactivity towards epitopes A, B, and C. Little reactivity was seen to epitope C, while as shown in Figures 15, 16, 17, and 18, epitopes A and B  
10 reacted with the majority of sera. However, antibodies to the most reactive epitope (epitope A) did not seem to predict remission of disease, while the anti-1bE1 antibodies (epitope B) were present almost exclusively in long term responders at the start of IFN treatment. Therefore, anti-1bE1 (epitope B) antibodies and anti-env53 (epitope A) antibodies could be shown to be useful markers for prognosis of hepatitis C disease. The env53 epitope may be  
15 advantageously used for the detection of cross-reactive antibodies (antibodies that cross-react between major genotypes) and antibodies to the env53 region may be very useful for universal E1 antigen detection in serum or liver tissue. Monoclonal antibodies that recognized the env53 region were reacted with a random epitope library. In 4 clones that reacted upon immunoscreening with the monoclonal antibody 5E1A10, the sequence -GWD-  
20 was present. Because of its analogy with the universal HCV sequence present in all HCV variants in the env53 region, the sequence AWD is thought to contain the essential sequence of the env53 cross-reactive murine epitope. The env31 clearly also contains a variable region which may contain an epitope in the amino terminal sequence -YQVRNSTGL- (SEQ ID NO:93) and may be useful for diagnosis. Env31 or E1-31 as shown in Table 3, is a part of the  
25 peptide 1bE1. Peptides E1-33 and E1-51 also reacted to some extent with the murine antibodies, and peptide E1-55 (containing the variable region 6 (V6); spanning amino acid positions 329-336) also reacted with some of the patient sera.

Anti-E2 antibodies clearly followed a different pattern than the anti-E1 antibodies, especially in patients with a long-term response to treatment. Therefore, it is clear that the  
30 decrease in anti-envelope antibodies could not be measured as efficiently with an assay employing a recombinant E1/E2 protein as with a single anti-E1 or anti-E2 protein. The anti-E2 response would clearly blur the anti-E1 response in an assay measuring both kinds of antibodies at the same time. Therefore, the ability to test anti-envelope antibodies to the

single E1 and E2 proteins, was shown to be useful.

#### 7.4. Mapping of anti-E2 antibodies

Of the 24 anti-E2 Mabs only three could be competed for reactivity to recombinant  
5 E2 by peptides, two of which reacted with the HVRI region (peptides E2-67 and E2-69,  
designated as epitope A) and one which recognized an epitope competed by peptide E2-13B  
(epitope C). The majority of murine antibodies recognized conformational anti-E2 epitopes  
(Figure 19). A human response to HVRI (epitope A), and to a lesser extent HVRII (epitope  
B) and a third linear epitope region (competed by peptides E2-23, E2-25 or E2-27,  
10 designated epitope E) and a fourth linear epitope region (competed by peptide E2-17B,  
epitope D) could also frequently be observed, but the majority of sera reacted with  
conformational epitopes (Figure 20). These conformational epitopes could be grouped  
according to their relative positions as follows: the IgG antibodies in the supernatant of  
hybridomas 15C8C1, 12D11F1, 9G3E6, 8G10D1H9, 10D3C4, 4H6B2, 17F2C2, 5H6A7,  
15 15B7A2 recognizing conformational epitopes were purified by means of protein A affinity  
chromatography and 1 mg/ml of the resulting IgG's were biotinylated in borate buffer in the  
presence of biotin. Biotinylated antibodies were separated from free biotin by means of  
gelfiltration chromatography. Pooled biotinylated antibody fractions were diluted 100 to  
10,000 times. E2 protein bound to the solid phase was detected by the biotinylated IgG in the  
20 presence of 100 times the amount of non-biotinylated competing antibody and subsequently  
detected by alkaline phosphatase labeled streptavidin.

Percentages of competition are given in Table 6. Based on these results, 4  
conformational anti-E2 epitope regions (epitopes F, G, H and I) could be delineated (Figure  
38). Alternatively, these Mabs may recognize mutant linear epitopes not represented by the  
25 peptides used in this study. Mabs 4H6B2 and 10D3C4 competed reactivity of 16A6E7, but  
unlike 16A6E7, they did not recognize peptide E2-13B. These Mabs may recognize variants  
of the same linear epitope (epitope C) or recognize a conformational epitope which is  
sterically hindered or changes conformation after binding of 16A6E7 to the E2-13B region  
(epitope H).

## **Example 8: E1 glycosylation mutants**

### **8.1. Introduction**

The E1 protein encoded by vvHCV10A, and the E2 protein encoded by vvHCV41 to  
5 44 expressed from mammalian cells contain 6 and 11 carbohydrate moieties, respectively.  
This could be shown by incubating the lysate of vvHCV10A-infected or vvHCV44-infected  
RK13 cells with decreasing concentrations of glycosidases (PNGase F or Endoglycosidase  
H, (Boehringer Mannheim Biochemica) according to the manufacturer's instructions), such  
that the proteins in the lysate (including E1) are partially deglycosylated (Fig. 39 and 40,  
10 respectively).

Mutants devoid of some of their glycosylation sites could allow the selection of  
envelope proteins with improved immunological reactivity. For HIV for example, gp120  
proteins lacking certain selected sugar-addition motifs, have been found to be particularly  
useful for diagnostic or vaccine purpose. The addition of a new oligosaccharide side chain in  
15 the hemagglutinin protein of an escape mutant of the A/Hong Kong/3/68 (H3N2) influenza  
virus prevents reactivity with a neutralizing monoclonal antibody (Skehel et al, 1984). When  
novel glycosylation sites were introduced into the influenza hemagglutinin protein by site-  
specific mutagenesis, dramatic antigenic changes were observed, suggesting that the  
carbohydrates serve as a modulator of antigenicity (Gallagher et al., 1988). In another  
20 analysis, the 8 carbohydrate-addition motifs of the surface protein gp70 of the Friend Murine  
Leukemia Virus were deleted. Although seven of the mutations did not affect virus  
infectivity, mutation of the fourth glycosylation signal with respect to the amino terminus  
resulted in a non-infectious phenotype (Kayman et al., 1991). Furthermore, it is known in the  
art that addition of N-linked carbohydrate chains is important for stabilization of folding  
25 intermediates and thus for efficient folding, prevention of misfolding and degradation in the  
endoplasmic reticulum, oligomerization, biological activity, and transport of glycoproteins  
(see reviews by Rose et al., 1988; Doms et al., 1993; Helenius, 1994).

After alignment of the different envelope protein sequences of HCV genotypes, it  
may be inferred that not all 6 glycosylation sites on the HCV subtype 1b E1 protein are  
30 required for proper folding and reactivity, since some are absent in certain (sub)types. The  
fourth carbohydrate motif (on Asn251), present in types 1b, 6a, 7, 8, and 9, is absent in all  
other types known today. This sugar-addition motif may be mutated to yield a type 1b E1  
protein with improved reactivity. Also the type 2b sequences show an extra glycosylation site

in the V5 region (on Asn299). The isolate S83, belonging to genotype 2c, even lacks the first carbohydrate motif in the V1 region (on Asn), while it is present on all other isolates (Stuyver et al., 1994). However, even among the completely conserved sugar-addition motifs, the presence of the carbohydrate may not be required for folding, but may have a role in evasion of immune surveillance. Therefore, identification of the carbohydrate addition motifs which are not required for proper folding (and reactivity) is not obvious, and each mutant has to be analyzed and tested for reactivity. Mutagenesis of a glycosylation motif (NXS or NXT sequences) can be achieved by either mutating the codons for N, S, or T, in such a way that these codons encode amino acids different from N in the case of N, and/or amino acids different from S or T in the case of S and in the case of T. Alternatively, the X position may be mutated into P, since it is known that NPS or NPT are not frequently modified with carbohydrates. After establishing which carbohydrate-addition motifs are required for folding and/or reactivity and which are not, combinations of such mutations may be made.

## 8.2. Mutagenesis of the E1 protein

All mutations were performed on the E1 sequence of clone HCC110A (SEQ ID NO:5). The first round of PCR was performed using sense primer 'GPT' (see Table 7) targetting the GPT sequence located upstream of the vaccinia 11K late promoter, and an antisense primer (designated GLY#, with # representing the number of the glycosylation site, see Fig. 41) containing the desired base change to obtain the mutagenesis. The six GLY# primers (each specific for a given glycosylation site) were designed such that:

- Modification of the codon encoding for the N-glycosylated Asn (AAC or AAT) to a Gln codon (CAA or CAG). Glutamine was chosen because it is very similar to asparagine (both amino acids are neutral and contain non-polar residues, glutamine has a longer side chain

- (one more -CH<sub>2</sub>- group).

- The introduction of silent mutations in one or several of the codons downstream of the glycosylation site, in order to create a new unique or rare (e.g. a second SmaI site for E1Gly5) restriction enzyme site. Without modifying the amino acid sequence, this mutation will provide a way to distinguish the mutated sequences from the original E1 sequence (pvHCV-10A) or from each other (Figure 41). This additional restriction site may also be useful for the construction of new hybrid (double, triple, etc.) glycosylation mutants.

- 18 nucleotides extend 5' of the first mismatched nucleotide and 12 to 16 nucleotides extend to the 3' end. Table 7 depicts the sequences of the six GLY# primers overlapping the

sequence of N-linked glycosylation sites.

For site-directed mutagenesis, the 'mispriming' or 'overlap extension' (Horton, 1993) was used. The concept is illustrated in Figures 42 and 43. First, two separate fragments were amplified from the target gene for each mutated site. The PCR product obtained from the 5' end (product GLY#) was amplified with the 5' sense GPT primer (see Table 7) and with the  
5 respective 3' antisense GLY# primers. The second fragment (product OVR#) was amplified with the 3' antisense TK<sub>R</sub> primer and the respective 5' sense primers (OVR# primers, see Table 7, Figure 43).

The OVR# primers target part of the GLY# primer sequence. Therefore, the two  
10 groups of PCR products share an overlap region of identical sequence. When these intermediate products are mixed (GLY-1 with OVR-1, GLY-2 with OVR-2, etc.), melted at high temperature, and reannealed, the top sense strand of product GLY# can anneal to the antisense strand of product OVR# (and vice versa) in such a way that the two strands act as primers for one another (see Fig. 42.B.). Extension of the annealed overlap by Taq  
15 polymerase during two PCR cycles created the full-length mutant molecule E1Gly#, which carries the mutation destroying the glycosylation site number #. Sufficient quantities of the E1GLY# products for cloning were generated in a third PCR by means of a common set of two internal nested primers. These two new primers are respectively overlapping the 3' end of the vaccinia 11K promoter (sense GPT-2 primer) and the 5' end of the vaccinia thymidine  
20 kinase locus (antisense TK<sub>R</sub>-2 primer, see Table 7). All PCR conditions were performed as described in Stuyver et al. (1993).

Each of these PCR products was cloned by EcoRI/BamHI cleavage into the EcoRI/BamHI-cut vaccinia vector containing the original E1 sequence (pvHCV-10A).

The selected clones were analyzed for length of insert by EcoRI/BamH I cleavage  
25 and for the presence of each new restriction site. The sequences overlapping the mutated sites were confirmed by double-stranded sequencing.

### 8.3. Analysis of E1 glycosylation mutants

Starting from the 6 plasmids containing the mutant E1 sequences as described in  
30 example 8.2, recombinant vaccinia viruses were generated by recombination with wt vaccinia virus as described in example 2.5. Briefly, 175 cm<sup>2</sup>-flasks of subconfluent RK13 cells were infected with the 6 recombinant vaccinia viruses carrying the mutant E1 sequences, as well as with the vvHCV-10A (carrying the non-mutated E1 sequence) and wt

vaccinia viruses. Cells were lysed after 24 hours of infection and analyzed on western blot as described in example 4 (see Figure 44A). All mutants showed a faster mobility (corresponding to a smaller molecular weight of approximately 2 to 3 kDa) on SDS-PAGE than the original E1 protein; confirming that one carbohydrate moiety was not added.

5 Recombinant viruses were also analyzed by PCR and restriction enzyme analysis to confirm the identity of the different mutants. Figure 44B shows that all mutants (as shown in Figure 41) contained the expected additional restriction sites. Another part of the cell lysate was used to test the reactivity of the different mutant by ELISA. The lysates were diluted 20 times and added to microwell plates coated with the lectin GNA as described in example 6.

10 Captured (mutant) E1 glycoproteins were left to react with 20-times diluted sera of 24 HCV-infected patients as described in example 6. Signal to noise (S/N) values (OD of GLY#/OD of wt) for the six mutants and E1 are shown in Table 8. The table also shows the ratios between S/N values of GLY# and E1 proteins. It should be understood that the approach to use cell lysates of the different mutants for comparison of reactivity with patient sera may

15 result in observations that are the consequence of different expression levels rather than reactivity levels. Such difficulties can be overcome by purification of the different mutants as described in example 5, and by testing identical quantities of all the different E1 proteins. However, the results shown in table 5 already indicate that removal of the 1st (GLY1), 3rd (GLY3), and 6th (GLY6) glycosylation motifs reduces reactivity of some sera, while removal

20 of the 2nd and 5th site does not. Removal of GLY4 seems to improve the reactivity of certain sera. These data indicate that different patients react differently to the glycosylation mutants of the present invention. Thus, such mutant E1 proteins may be useful for the diagnosis (screening, confirmation, prognosis, etc.) and prevention of HCV disease.

25 **Example 9: Expression of HCV E2 protein in glycosylation-deficient yeasts**

The E2 sequence corresponding to clone HCCl41 was provided with the  $\alpha$ -mating factor pre/pro signal sequence, inserted in a yeast expression vector and *S. cerevisiae* cells transformed with this construct secreted E2 protein into the growth medium. It was observed that most glycosylation sites were modified with high-mannose type glycosylations upon

30 expression of such a construct in *S. cerevisiae* strains (Figure 45). This resulted in a too high level of heterogeneity and in shielding of reactivity, which is not desirable for either vaccine or diagnostic purposes. To overcome this problem, *S. cerevisiae* mutants with modified glycosylation pathways were generated by means of selection of vanadate-resistant clones.

Such clones were analyzed for modified glycosylation pathways by analysis of the molecular weight and heterogeneity of the glycoprotein invertase. This allowed us to identify different glycosylation deficient *S. cerevisiae* mutants. The E2 protein was subsequently expressed in some of the selected mutants and left to react with a monoclonal antibody as described in example 7, on western blot as described in example 4 (Figure 46).

#### **Example 10. General utility**

The present results show that not only a good expression system but also a good purification protocol are required to reach a high reactivity of the HCV envelope proteins with human patient sera. This can be obtained using the proper HCV envelope protein expression system and/or purification protocols of the present invention which guarantee the conservation of the natural folding of the protein and the purification protocols of the present invention which guarantee the elimination of contaminating proteins and which preserve the conformation, and thus the reactivity of the HCV envelope proteins. The amounts of purified HCV envelope protein needed for diagnostic screening assays are in the range of grams per year. For vaccine purposes, even higher amounts of envelope protein would be needed. Therefore, the vaccinia virus system may be used for selecting the best expression constructs and for limited upscaling, and large-scale expression and purification of single or specific oligomeric envelope proteins containing high-mannose carbohydrates may be achieved when expressed from several yeast strains. In the case of hepatitis B for example, manufacturing of HBsAg from mammalian cells was much more costly compared with yeast-derived hepatitis B vaccines.

The purification method disclosed in the present invention may also be used for 'viral envelope proteins' in general. Examples are those derived from Flaviviruses, the newly discovered GB-A, GB-B and GB-C Hepatitis viruses, Pestiviruses (such as Bovine viral Diarrhoea Virus (BVDV), Hog Cholera Virus (HCV), Border Disease Virus (BDV)), but also less related viruses such as Hepatitis B Virus (mainly for the purification of HBsAg).

The envelope protein purification method of the present invention may be used for intra- as well as extracellularly expressed proteins in lower or higher eukaryotic cells or in prokaryotes as set out in the detailed description section.

#### **Example 11. Demonstration of Prophylactic and Therapeutic Utility**

Liver disease in chimpanzees chronically infected with HCV can be reduced by



immunization with E1. Multiple immunizations, however, were required in order to reach a significant immune response. One of ordinary skill will appreciate that viral persistence is produced with immune modulation which is either orchestrated by the virus itself or by the host. In order to analyze if such an immune modulation does exist in HCV, the immune responses against E1 and NS3 in naive and chronically infected chimpanzees were compared. Since a lower response in the chronically infected animals was anticipated, this group of animals was selected for a more rigorous immunization schedule including the following: use of an adjuvant proven in mice to be more potent for inducing cellular responses (Table 9) compared to alum, which was the adjuvant used for naive animals; and the immunization schedule for chronically infected animals consisted of 12 immunizations compared to 6 for naive animals (Fig. 47).

Although the number of immunized animals does not allow statistical analysis, the following clear tendency can be detected in the humoral responses (Table 10): the number of immunizations for seroconversion is lower in naive animals; and the magnitude of the immune response is substantially greater in the naive animals, 2/3 infected animals do not reach the level of 10 internal units, even after 12 immunizations.

The analysis of the cellular responses, after three immunizations, reveals an even larger difference (Fig. 48a-d), including the following: E1-specific T-cell proliferation is almost absent in the chronically infected animals, while a clear stimulation can be seen in the naive setting; IL-2 measurements confirmed that the low stimulation of the T-cell compartment in chronic carriers; and, a clear Th2 (IL-4) response in naive animals is induced, as expected for an alum-adjuvant containing vaccine.

This confirms that at least E1 immunization provides a prophylactic effect in naive animals and suggest that E2 and/or combinations of E1 and E2 proteins and/or peptides may provide useful therapeutic and/or prophylactic benefits in naive animals.

The 'impairment' to induce both cellular and humoral responses against an HCV E1 antigen can be only partially overcome by multiple immunizations, as demonstrated by the following results: an increase in antibody titer after each injection was noted but the levels as in naive animals were not reached in 2/3 animals; and the T-cell proliferative responses remain very low (Fig. 49). The ELISPOT results show, however, a minor increase in IL-2 (not shown), no change in IFN-g (not shown) and an increase in IL-4 (Fig. 49) which indicates that Th2 type responses are more readily induced. IL-4 was noted to remain at a low level compared to the level reached after three immunizations in naive animals.

A quite similar observation was made for NS3 immunizations where an even stronger adjuvant (RIBI) was used in the chronic chimpanzee. As compared with an alum formulation in naive animals the following has been noted: the induced antibody titers are comparable in both groups (not shown); and both cytokine secretion and T-cell proliferation are almost absent in the chronic animals compared to the responses in naive animals (Fig. 49a-b).

Currently there have been some indications that immune responses against HCV in chronic carriers are low or at least insufficient to allow clearance of infection. The above results support the hypothesis that the immune system of HCV chronic carriers may be impaired and that they do not respond to HCV antigens as efficiently as in a naive situation.

In a study by Wiedmann et al., (Hepatology 2000; 31: 230-234), vaccination for HBV was less effective in HCV chronic carriers, which indicates that such an immune impairment is not limited to HCV antigens. De Maria et al. (Hepatology 2000; 32: 444-445) confirmed these data and have proposed adapted vaccine dosing regimens for HCV patients. The data presented herein indicates that increasing the number of immunizations may indeed augment humoral responses but that cellular (especially Th1) responses are difficult to induce, even when powerful adjuvants are used. It may be advantageous to begin immunization at the time of antiviral therapy, when the immune system is more prone to respond.

**Table 1: Recombinant vaccinia plasmids and viruses**

Plasmid Name	HCV cDNA subclone			Vector used for insertion
	Name	Construction	Length (nt/aa)	
pvHCV-13A	E1s	EcoR I - Hind III	472/157	pgptATA-18
pvHCV-12A	E1s	EcoR I - Hind III	472/158	pgptATA-18
pvHCV-9A	E1	EcoR I - Hind III	631/211	pgptATA-18
pvHCV-11A	E1s	EcoR I - Hind III	625/207	pgptATA-18
pvHCV-17A	E1s	EcoR I - Hind III	625/208	pgptATA-18
pvHCV-10A	E1	EcoR I - Hind III	783/262	pgptATA-18
pvHCV-18A	COREs	Acc I (Kl) - EcoR I (Kl)	403/130	pgptATA-18
pvHCV-34	CORE	Acc I (Kl) - Fsp I	595/197	pgptATA-18
pvHCV-33	CORE-E1	Acc I (Kl)	1150/380	pgptATA-18
pvHCV-35	CORE-E1b.his	EcoR I - BamH I (Kl)	1032/352	pMS-66
pvHCV-36	CORE-E1n.his	EcoR I - Nco I (Kl)	1106/376	pMS-66
pvHCV-37	E1 $\Delta$	Xma I - BamH I	711/239	pvHCV-10A
pvHCV-38	E1 $\Delta$ s	EcoR I - BstE II	553/183	pvHCV-11A
pvHCV-39	E1 $\Delta$ b	EcoR I - BamH I	960/313	pgsATA-18
pvHCV-40	E1 $\Delta$ b.his	EcoR I - BamH I (Kl)	960/323	pMS-66
pvHCV-41	E2bs	BamH I (Kl)-AlwN I (T4)	1005/331	pgsATA-18
pvHCV-42	E2bs.his	BamH I (Kl)-AlwN I (T4)	1005/341	pMS-66
pvHCV-43	E2ns	Nco I (Kl) - AlwN I (T4)	932/314	pgsATA-18
pvHCV-44	E2ns.his	Nco I (Kl) - AlwN I (T4)	932/321	pMS-66
pvHCV-62	E1s (type 3a)	EcoR I - Hind III	625/207	pgsATA-18
pvHCV-63	E1s (type 5)	EcoR I - Hind III	625/207	pgsATA-18
pvHCV-64	E2	BamH I - Hind III	1410/463	pgsATA-18
pvHCV-65	E1-E2	BamH I - Hind III	2072/691	pvHCV-10A
pvHCV-66	CORE-E1-E2	BamH I - Hind III	2427/809	pvHCV-33

nt: nucleotide; aa: amino acid; Kl: Klenow DNA Pol filling; T4: T4 DNA Pol filling

Position: amino acid position in the HCV polyprotein sequence

**Table 1 - continued: Recombinant vaccinia plasmids and viruses**

Plasmid Name	HCV cDNA subclone			Vector used for insertion
	Name	Construction	Length (nt/aa)	
pvHCV-81	E1*-GLY 1	EcoRI – BamH I	783/262	pvHCV-10A
pvHCV-82	E1*-GLY 2	EcoRI – BamH I	783/262	pvHCV-10A
pvHCV-83	E1*-GLY 3	EcoRI – BamH I	783/262	pvHCV-10A
pvHCV-84	E1*-GLY 4	EcoRI – BamH I	783/262	pvHCV-10A
pvHCV-85	E1*-GLY 5	EcoRI – BamH I	783/262	pvHCV-10A
pvHCV-86	E1*-GLY 6	EcoRI – BamH I	783/262	pvHCV-10A

nt: nucleotide; aa: amino acid; KI: Klenow DNA Pol filling; T4: T4 DNA Pol filling

Position: amino acid position in the HCV polyprotein sequence

**Table 2 : Summary of anti-E1 tests****S/N  $\pm$  SD (mean anti-E1 titer)**

	Start of treatment	End of treatment	Follow-up
LTR	6.94 $\pm$ 2.29 (1:3946)	4.48 $\pm$ 2.69 (1:568)	2.99 $\pm$ 2.69 (1:175)
NR	5.77 $\pm$ 3.77 (1:1607)	5.29 $\pm$ 3.99 (1:1060)	6.08 $\pm$ 3.73 (1:1978)

**LTR** : Long-term, sustained response for more than 1 year**NR** : No response, response with relapse, or partial response

**Table 3 : Synthetic peptides for competition studies**

PROTEIN	PEPTIDE	AMINO ACID SEQUENCE	POSITION	SEQ ID NO
E1	E1-31	LLSCLTVPASAYQVRNSTGL	181-200	56
	E1-33	QVRNSTGLYHVTNDCPNSSI	193-212	57
	E1-35	NDCPNSSIVYEAHDAILHTP	205-224	58
	E1-35A	SNSSIVYEAADMIMHTPGCV	208-227	59
	E1-37	HDAILHTPGCVPCVREGNVS	217-236	60
	E1-39	CVREGNVSRCWVAMTPTVAT	229-248	61
	E1-41	AMTPTVATRDGKLPATQLRR	241-260	62
	E1-43	LPATQLRRHIDLLVGSATLC	253-272	63
	E1-45	LVGSATLCSALYVGDLCSV	265-284	64
	E1-49	QLFTFSPRRHWTQGCNC SI	289-308	65
	E1-51	TQGCNC SIYPGHITGHRMAW	301-320	66
	E1-53	ITGHRMAWDMMNWSPTAAL	313-332	67
	E1-55	NWSPTAALVMAQLLRIPQAI	325-344	68
	E1-57	LLRIPQAILDMIAGAHWGV L	337-356	69
	E1-59	AGAHWGVLAGIAYFSMVGNM	349-368	70
	E1-63	VVLLLFAGVDAETIVSGGQA	373-392	71
E2	E2-67	SGLVSLFTPGAKQNIQLINT	397-416	72
	E2-69	QNIQLINTNGSWHINSTALN	409-428	73
	E2-33B	LNCNESLNTGWLAGLIYQHK	427-446	74
	E2-31B	AGLIYQHKFNSSGCPERLAS	439-458	75
	E2-1B	GCPERLASCRPLTDFDQGWG	451-470	76
	E2-3B	TDFDQGWGPISYANGSGPDQ	463-482	77

**Table 3 - continued : Synthetic peptides for competition studies**

E2-5B	ANGSGPDQRPYCWHYPPKPC	475-494	78
E2-7B	WHYPPKPCGIVPAKSVCGPV	487-506	79
E2-9B	AKSVCGPVYCFTPSPVVVGT	499-518	80
E2-11B	PSPVVVGTTDRSGAPTYSWG	511-530	81
E2-13B	GAPTYSWGENDTDVFLNNT	523-542	82
E2-17B	GNWFGCTWMNSTGFTKVCGA	547-566	83
E2-19B	GFTKVCGAPPVCIGGAGNNT	559-578	84
E2-21	IGGAGNNTLHCPTDCFRKHP	571-590	85
E2-23	TDCFRKHPDATYSRCGSGPW	583-602	86
E2-25	SRCGSGPWITPRCLVDYPYR	595-614	87
E2-27	CLVDYPYRLWHYPCTINYTI	607-626	88
E2-29	PCTINYTIFKIRMYVGGVEH	619-638	89
E2-31	MYVGGVEHRLEAACNWPGE	631-650	90
E2-33	ACNWPGERCDLEDNRSEL	643-662	91
E2-35	EDRDRSELSPLLLTTTQWQV	655-674	92

**Table 4. Change of Envelope Antibody levels over time (complete study, 28 patients)**

Wilcoxon Signed Rank test (P values)	E1Ab NR	E1Ab NR All	E1Ab NR	E1Ab LTR type 1b	E1Ab LTR type 3a	E1Ab LTR All	E2Ab NR	E1Ab LTR		
								type 1b	type 3a	All
End of therapy*	0.1167	0.2604	0.285		0.0058**	0.043**	0.0499**	0.0186**	0.0640	
6 months follow up*	0.86		0.7213	0.5930	0.0047**	0.043**	0.063		0.04326	0.0464**
12 months follow up*	0.7989	0.3105	1		0.0051**	0.0679	0.0277**	0.0869	0.0058**	

\*Data were compared with values obtained at initiation of therapy

\*\*P values &lt; 0.05



**Table 5. Difference between LTR and NR (complete study)**

Mann-Withney	E1Ab S/N	E1Ab titers	E1Ab S/N type 1b type 3aAll	E1Ab S/N	E2Ab S/N
U test (P values)	All	All			
Initiation of therapy	0.0257*	0.05*	0.68	0.1078	
End of therapy	0.1742			0.1295	
6 months follow up	1	0.6099	0.425	0.3081	
12 months follow up	0.67	0.23	0.4386	0.6629	

\* P values &lt; 0.05



Table 7. Primers

	SEQ ID NO: 96	GPT	5'-GTTTAACCACTGCATGATG-3'
5	SEQ ID NO: 97	TK <sub>R</sub>	5'-GTCCCATCGAGTGGGCTAC-3'
	SEQ ID NO: 98	GLY1	5'-CGTGACATGGTACAT <u>TCCGGACACTTGGCGCACTTCATAAGCGGA</u> -3'
	SEQ ID NO: 99	GLY2	5'-TGCCTCATACACAATG <u>GAGCTCTCTGGGACGAGTCGTTTCGTGAC</u> -3'
	SEQ ID NO: 100	GLY3	5'-TACCCAGCAGCGGGAG <u>GCTCTGTGCTCCGAAACGACGGGCAC</u> -3'
	SEQ ID NO: 101	GLY4	5'-TGTCGTGGTGGGACG <u>GAGGCCCTGCC</u> TAGCTGGAGCGTGGG-3'
10	SEQ ID NO: 102	GLY5	5'-CGTTATGTGGCCCGGGTAGATTGAGCAC <u>TGGCAGTCC</u> TGCACCGTCTC-3'
	SEQ ID NO: 103	GLY6	5'-CAGGGCCGTTGTAGGC <u>CTCCA</u> CTGCATCATCATATCCCAAGC-3'
	SEQ ID NO: 104	OVR1	5'-CCGGAATGTACCATGTACGAACGAC-3'
	SEQ ID NO: 105	OVR2	5'- <u>GCTCC</u> ATTGTGTATGAGGCAGCGG-3'
	SEQ ID NO: 106	OVR3	5'- <u>GAGCT</u> CCCCGCTGCTGGGTAGCGC-3'
15	SEQ ID NO: 107	OVR4	5'-CC <u>TCC</u> GTCCCCACCAACGACAAATACG-3'
	SEQ ID NO: 108	OVR5	5'-CTA <u>CCCCGGG</u> CCACATAACGGGTCACCG-3'
	SEQ ID NO: 109	OVR6	5'-GG <u>AGGCC</u> TACAACGGCCCTGGTGG-3'
	SEQ ID NO: 110	GPT-2	5'-TTCTATCGATTAAATAGAATC -3'
20	SEQ ID NO: 111	TK <sub>R</sub> -2	5'-GCCATACGCTCACAGCCGATCCCC-3'

20 nucleotides underlined represent additional restriction site

nucleotides in bold represent mutations with respect to the original HCC110A sequence

Table 8. Analysis of E1 glycosylation mutants by ELISA

SERUM										
	1	2	3	4	5	6	7	8	9	
SN GLY1	1.802462	2.120971	1.403871	1.205597	2.120191	2.866913	1.950345	1.866183	1.730193	
SN GLY2	2.400795	1.76818	2.325495	2.639308	2.459019	5.043993	2.146302	1.595477	1.688973	
SN GLY3	1.642718	1.715477	2.261646	2.354748	1.591818	4.833742	1.96692	1.482099	1.602222	
SN GLY4	2.578154	3.824038	3.874605	1.499387	3.15	4.71302	4.198751	3.959542	3.710507	
SN GLY5	2.482051	1.793761	2.409344	2.627358	1.715311	4.964765	2.13912	1.576336	1.708937	
SN GLY6	2.031487	1.495737	2.131613	2.527925	2.494833	4.784027	2.02069	1.496489	1.704976	
SN E1	2.828205	2.227036	2.512792	2.790881	3.131579	4.869128	2.287753	1.954198	1.805556	

	10	11	12	13	14	15	16	17	18	
SN GLY1	2.468162	1.220654	1.629403	5.685561	3.233604	3.763498	1.985105	2.317721	6.675179	
SN GLY2	2.482212	1.467582	2.070524	7.556682	2.567613	3.621928	3.055649	2.933792	7.65433	
SN GLY3	2.191558	1.464216	1.721164	7.930538	2.763055	3.016099	2.945628	2.515305	5.775357	
SN GLY4	5.170841	4.250784	3.955153	8.176816	6.561122	5.707668	5.684498	5.604813	6.4125	
SN GLY5	3.021807	1.562092	2.07278	8.883408	2.940334	3.125561	3.338912	2.654224	5.424107	
SN GLY6	2.677757	1.529608	1.744221	8.005561	2.499952	2.621704	2.572385	2.363301	5.194107	
SN E1	2.616822	1.55719	2.593886	8.825112	3.183771	3.067265	3.280335	2.980354	7.191964	

SUM AVERAGE									
	19	20	21	22	23	24	S/N	S/N	S/N
SN GLY1	1.93476	2.47171	4.378633	1.188748	2.158889	1.706992	59.88534	2.495223	
SN GLY2	2.127712	2.921288	4.680101	1.150781	1.661914	1.632785	69.65243	2.902185	
SN GLY3	1.980185	2.557384	4.268633	0.97767	1.336775	1.20376	62.09872	2.587447	
SN GLY4	3.813321	3.002535	4.293038	2.393011	3.68213	2.481585	102.6978	4.279076	
SN GLY5	2.442804	3.126761	4.64557	1.153656	1.817901	1.638211	69.26511	2.886046	
SN GLY6	1.506716	2.665433	2.781063	1.280743	1.475062	1.716423	61.32181	2.555075	
SN E1	2.771218	3.678068	5.35443	1.167286	2.083333	1.78252	76.54068	3.189195	

Table 8 (continued). Analysis of E1 glycosylation mutants by ELISA

**SERUM**

	1	2	3	4	5	6	7	8	9
GLY1/E1	0.637316	0.952374	0.55869	0.431977	0.677036	0.588794	0.852516	0.954961	0.958261
GLY2/E1	0.848876	0.793961	0.925463	0.94569	0.785233	1.035913	0.93817	0.816436	0.935431
GLY3/E1	0.580834	0.770296	0.900053	0.84373	0.508312	0.992733	0.859761	0.758418	0.887385
GLY4/E1	0.911587	1.717097	1.541952	0.537245	1.005882	0.967939	1.835317	2.026172	2.05505
GLY5/E1	0.877607	0.805447	0.958831	0.941408	0.547746	1.019642	0.935031	0.806641	0.946488
GLY6/E1	0.718296	0.671626	0.848305	0.90578	0.796669	0.982522	0.883264	0.765781	0.944294

	10	11	12	13	14	15	16	17	18
GLY1/E1	0.94319	0.783882	0.628171	0.644248	1.015652	1.226988	0.605153	0.777666	0.928144
GLY2/E1	0.94856	0.942455	0.798232	0.85627	0.806469	1.180833	0.931505	0.984377	1.064289
GLY3/E1	0.837488	0.940294	0.663547	0.898633	0.867856	0.983319	0.897966	0.843962	0.803029
GLY4/E1	1.976	2.72978	1.524798	0.92654	2.060802	1.860833	1.732902	1.880587	0.89162
GLY5/E1	1.154762	1.003148	0.799102	1.006606	0.923538	1.019006	1.017857	0.890574	0.75419
GLY6/E1	1.023286	0.982288	0.672435	0.907134	0.785217	0.854737	0.784184	0.79296	0.72221

	19	20	21	22	23	24	SUM	AVERAGE
GLY1/E1	0.698162	0.672013	0.817759	1.018386	1.036267	0.957628	E1/GLY# 19.36524	E1/GLY# 0.806885
GLY2/E1	0.76779	0.794245	0.874061	0.98586	0.797719	0.915998	21.67384	0.903077
GLY3/E1	0.714554	0.695306	0.797215	0.837558	0.641652	0.675314	19.19921	0.799967
GLY4/E1	1.376045	0.816335	0.801773	2.050064	1.767422	1.392178	36.38592	1.51608
GLY5/E1	0.881491	0.850109	0.867612	0.988323	0.872593	0.919042	21.78679	0.907783
GLY6/E1	0.543702	0.724683	0.519395	1.097197	0.70803	0.962919	19.59691	0.816538

**Table 9. Profile of adjuvated E1 in Balb/c mice**

	alum	T-cell adjuvant	RBI
antibody titre (mean $\pm$ SD, n=6)	96000 $\pm$ 101000	62000 $\pm$ 60000	176000 $\pm$ 149000
antibody isotypes	IgG1	IgG1/2b	IgG1/2a
T-cell prelieration in spleen <sup>1</sup> (n=3)	11750 (2/3)	48300 (3/3)	26000 (3/3)
T-cell proliferation in lymph node <sup>2</sup>	no specific stimulation	4000	8000
cytokine profile (spleen)	Il-4	IFN-g/Il-4	IFN-g/Il-4

5   <sup>1</sup> after three s.c./i.m. immunizations, 3 randomly selected mice were analyzed individually, the result is expressed as the mean specific cpm obtained after 4 days of E1 stimulation (1  $\mu$ g/ml), the number in brackets refers to the number of mice with specific stimulation above background

10   <sup>2</sup> after one single intra footpath immunization (n=2), the result is expressed as the mean specific cpm obtained after 5 days of E1 stimulation (1  $\mu$ g/ml)

**Table 10. Humoral Responses: No. of immunizations required for different E-1 antibodies levels**

Animal	status	seroconversion <sup>1</sup>	> 1 U/ml <sup>2</sup>	> 10 U/ml
Marcel	chronic	3	4	5
Peggy	chronic	3	5	>12
Femma	chronic	4	5	>12
Yoran	naive	3	4	5
Marti	naive	2	3	5

5

<sup>1</sup> defined as ELISA signal higher than cut-off level if no E1-antibodies were present prior to immunization, in the other cases the observation of a titer higher than the 3 individual time points of pre-immunization titers was considered as the point of seroconversion.

10 <sup>2</sup> the unit is defined as follows: the level of E1 antibodies in human chronic carriers prior to interferon therapy and infected with genotype 1b is < 0.1 U/ml for 50% of the patients, between 0.1 to 1 U/ml for 25% of the patients and > 1 U/ml in the remaining 25% of patients, n=58

**Example 12: Immunization of a chimpanzee chronically infected with HCV subtype 1b**

A chimpanzee (Phil) already infected for over 13 years (5015 days before immunization) with an HCV subtype 1b strain was vaccinated with E1 (aa 192-326) which was derived from a different strain of genotype 1 b, with a 95.1% identity on the amino acid level (see also Table 2 of WO 99/67285 the whole of which is incorporated herein by reference), and which was prepared as described in examples 1-3 of WO 99/97285. The chimpanzee received in total 6 intramuscular immunizations of each 50 µg E1 in PBS/0.05% CHAPS mixed with RIBI R-730 (MPLA+TDM+CWS) according to the manufacturer's protocol (Ribi Inc. Hamilton, MT). The 6 immunizations were given in two series of three shots with a three week interval and with a lag period of 6 weeks between the two series. Starting 150 days prior to immunization, during the immunization period and until 1 year post immunization (but see below and WO 99/67285) the chimpanzee was continuously monitored for various parameters indicative for the activity of the HCV induced disease. These parameters included blood chemistry, ALT, AST, gammaGT, blood chemistry, viral load in the serum, viral load in the liver and liver histology. In addition, the immune answer to the immunization was monitored both on the humoral and cellular level. During this period the animal was also monitored for any adverse effects of the immunization, such as change in behaviour, clinical symptoms, body weight, temperature and local reactions (redness, swelling, indurations). Such effects were not detected.

Clearly, ALT (and especially gammaGT, data not shown) levels decreased as soon as the antibody level against E1 reached its maximum (see, Figure 8 of WO 99/67285). ALT rebounded rather rapidly as soon as the antibody levels started to decline, but gammaGT remained at a lower level as long as anti-E1 remained detectable.

E2 antigen in the liver decreased to almost undetectable levels during the period in which anti-E1 was detectable and the E2 antigen rebounded shortly after the disappearance of these antibodies. Together with the Core and E2 antigen becoming undetectable in the liver, the inflammation of the liver markedly decreased (see also Table 3 of WO 99/67285). This is a major proof that the vaccine induces a reduction of the liver damage, probably by clearing, at least partially, the viral antigens from its major target organ, the liver.

The viraemia level, as measured by Amplicor HCV Monitor (Roche, Basel, Switzerland),



remained approximately unchanged in the serum during the whole study period.

More detailed analyses of the humoral response revealed that the maximum end-point titer reached  $14.5 \times 10^3$  (after the sixth immunization) and that this titer dropped to undetectable  
5 1 year post immunization (Figure 8 of WO 99/67285). Figure 9 of WO 99/67285 shows that the main epitopes, which can be mimicked by peptides, recognized by the B-cells are located at the N-terminal region of E2 (peptides V1V2 and V2V3, for details on the peptides used see Table 4 of WO 99/67285). Since the reactivity against the recombinant E1 is higher and longer lasting, it can also be deduced from this figure, that the antibodies  
10 recognizing these peptides represent only part of the total antibody population against E1. The remaining part is directed against epitopes which cannot be mimicked by peptides, i.e discontinuous epitopes. Such epitopes are only present on the complete E1 molecule or even only on the particle-like structure. Such an immune response against E1 is unique, at least compared to what is normally observed in human chronic HCV carriers  
15 (WO 96/13590 to Maertens et al.) and in chimpanzees (van Doorn et al., 1996), who raise anti-E1 antibodies in their natural course of infection. In those patients, anti-E1 is in part also directed to discontinuous epitopes but a large proportion is directed against the C4 epitope ( $\pm 50\%$  of the patient sera), a minor proportion against V1V2 (ranging from 2-70% depending on the genotype), and reactivity against V2V3 was only exceptionally recorded  
20 (Maertens et al., 1997).

Analysis of the T-cell reactivity indicated that also this compartment of the immune system is stimulated by the vaccine in a specific way, as the stimulation index of these T-cells rises from 1 to 2.5, and remains somewhat elevated during the follow up period (Figure 10 of  
25 WO 99/67285). It is this T cell reactivity that is only seen in Long term responders to interferon therapy (see: PCT/EP 94/03555 to Leroux-Roels et al.; Leroux-Roels et al., 1996).

**Example 13: Immunization of a chronic HCV carrier with different subtype**

30 A chimpanzee (Ton) already infected for over 10 years (3809 days before immunization) with HCV from genotype 1a was vaccinated with E1 from genotype 1b, with only a 79.3 % identity on the amino acid level (see also Table 2 of WO 99/67285), and prepared as described in the previous examples. The chimpanzee received a total of 6 intramuscular immunizations of 50  $\mu$ g E1 in PBS/0.05% CHAPS each mixed with RIBI R-730 according

to the manufacturer's protocol (Ribi Inc. Hamilton, MT). The 6 immunizations were given in two series of three shots with a three week interval and with a lag period of 4 weeks between the two series. Starting 250 days prior to immunization, during the immunization period and until 9 months (but see below and WO 99/67285) post immunization the chimpanzee was continuously monitored for various parameters indicative for the activity of the HCV induced disease. These parameters included blood chemistry, ALT, AST, gammaGT, viral load in the serum, viral load in the liver and liver histology. In addition, the immune answer to the immunization was monitored both on the humoral and cellular level. During this period the animal was also monitored for any adverse effects of the immunization, such as change in behaviour, clinical symptoms, body weight, temperature and local reactions (redness, swelling, indurations). Such effects were not detected.

Clearly, ALT levels (and gammaGT levels, data not shown) decreased as soon as the antibody level against E1 reached its maximum (Figure 11 of WO 99/67285). ALT and gammaGT rebounded as soon as the antibody levels started to decline, but ALT and gammaGT remained at a lower level during the complete follow up period. ALT levels were even significantly reduced after vaccination ( $62 \pm 6$  U/l) as compared to the period before vaccination ( $85 \pm 11$  U/l). Since less markers of tissue damage were recovered in the serum, these findings were a first indication that the vaccination induced an improvement of the liver disease.

E2 antigen levels became undetectable in the period in which anti-E1 remained above a titer of  $1.0 \times 10^3$ , but became detectable again at the time of lower E1 antibody levels. Together with the disappearance of HCV antigens, the inflammation of the liver markedly decreased from moderate chronic active hepatitis to minimal forms of chronic persistent hepatitis (Table 3 of WO 99/67285). This is another major proof that the vaccine induces a reduction of the liver damage, probably by clearing, at least partially, the virus from its major target organ, the liver.

The viraemia level, as measured by Amplicor HCV Monitor (Roche, Basel, Switzerland), in the serum remained at approximately similar levels during the whole study period. More detailed analysis of the humoral response revealed that the maximum end-point titer reached was  $30 \times 10^3$  (after the sixth immunization) and that this titer dropped to  $0.5 \times 10^3$  nine months after immunization (Figure 11 of WO 99/67285). Figure 12 of WO 99/67285

shows that the main epitopes, which can be mimicked by peptides and are recognized by the B-cells, are located at the N-terminal region (peptides V1V2 and V2V3, for details on the peptides used see Table 4 of WO 99/67285). Since the reactivity against the recombinant EI is higher and longer lasting, it can also be deduced from this figure, that the antibodies recognizing these peptides represent only part of the total antibody population against EI. The remaining part is most likely directed against epitopes which cannot be mimicked by peptides, i.e. discontinuous epitopes. Such epitopes are probably only present on the complete EI molecule or even only on the particle-like structure. Such an immune response against EI is unique, at least compared to what is normally observed in human chronic HCV carriers, which have detectable anti-EI. In those patients, anti-EI is in part also discontinuous, but a large proportion is directed against the C4 epitope (50% of the patient sera), a minor proportion against V1V2 (ranging from 2-70% depending on the genotype) and exceptionally reactivity against V2V3 was recorded (Maertens et al., 1997). As this chimpanzee is infected with an 1a isolate the antibody response was also evaluated for cross-reactivity towards a EI-1a antigen. As can be seen in Figure 13 of WO 99/67285, such cross-reactive antibodies are indeed generated, although, they form only part of the total antibody population. Remarkable is the correlation between the reappearance of viral antigen in the liver and the disappearance of detectable anti-1a EI antibodies in the serum.

20

Analysis of the T-cell reactivity indicated that also this compartment of the immune system is stimulated by the vaccine in a specific way, as the stimulation index of these T-cells rises from 0.5 to 5, and remains elevated during the follow up period (Figure 14 of WO 99/67285).

25

#### **Example 14: Reboosting of HCV chronic carriers with EI**

As the EI antibody titers as observed in examples 12 and 13 were not stable and declined over time, even to undetectable levels for the 1b infected chimp, it was investigated if this antibody response could be increased again by additional boosting. Both chimpanzees were immunized again with three consecutive intramuscular immunization with a three week interval (50 µg EI mixed with RIBI adjuvant). As can be judged from Figures 8 and 11 of WO 99/67285, the anti-EI response could indeed be boosted, once again the viral antigen in the liver decreased below detection limit. The viral load in the serum remained constant

30

although in Ton (Figure 11 of WO 99/67285). A viremia level of  $< 10^5$  genome equivalents per ml was measured for the first time during the follow up period.

Notable is the finding that, as was already the case for the first series of immunizations, the chimpanzee infected with the subtype 1b HCV strain (Phil) responds with lower anti-EI titers, than the chimpanzee infected with subtype 1a HCV strain (maximum titer in the first round  $14.5 \times 10^3$  versus  $30 \times 10^3$  for Ton and after additional boosting only  $1.2 \times 10^3$  for Phil versus  $40 \times 10^3$  for Ton). Although for both animals the beneficial effect seems to be similar, it could be concluded from this experiment that immunization of a chronic carrier with an EI protein derived from another subtype or genotype may be especially beneficial to reach higher titers, maybe circumventing a preexisting and specific immune suppression existing in the host and induced by the infecting subtype or genotype. Alternatively, the lower titers observed in the homologous setting (1b vaccine +1b infection) may indicate binding of the bulk of the antibodies to virus. Therefore, the induced antibodies may possess neutralizing capacity .

**Example 15: Demonstration of prophylactic utility of EI-vaccination in chimpanzee**

The HCV EI protein (amino acids 192-326) was expressed in Vero cells using recombinant vaccinia virus HCV11B. This vaccinia virus is essentially identical to vvHCV11A (as described in U.S. Patent No. 6,150,134, the entire contents of which is hereby incorporated by reference) but has been passaged from RK13 to Vero cells. The protein was purified (by means of lentil chromatography, reduction-alkylation and size exclusion chromatography) essentially as described in example 9 of PCT/E99/04342 (WO 99/67285), making use of iodoacetamide as alkylating agent for the cysteines. After purification the 3% empigen-BB was exchanged to 3% betain by size exclusion chromatography as described in example 1 of PCT/E99/04342 this process allows to recover EIs as a particle. Finally the material was desalted to PBS containing 0.5% betain and an EI concentration of 500  $\mu\text{g/ml}$ . This EI was mixed with an equal volume of Alhydrogel 1.3% (Superfos, Denmark) and finally further diluted with 8 volumes of 0.9% NaCl to yield alum-adjuvanted EI at a concentration of 50  $\mu\text{g EI/ml}$  and 0.13% of Alhydrogel.

The HCV E2deltaHVRI (amino acids 412-715) was expressed in and purified from Vero essentially as described for EI using recombinant vaccinia virus HCV101 which has been recombined from pvHCV-101 described in Example 8 of PCT/E99/04342 and wild type vaccinia virus. Also E2deltaHVRI behaves as a particle (measured by dynamic light scattering) after exchange of empigen to betain.

Five chimpanzees were selected which tested negative for HCV-RNA and HCV-antibodies. One of the animals (Huub) was not immunized, 2 animals received 6 immunizations with 50 µg EI adjuvanted with alum (Marti and Yoran) while the remaining 2 animals received 6 immunizations with 50 µg E2deltaHVRI adjuvanted with alum (Joost and Karlien). All immunizations were administered intra-muscularly with a 3 week interval. Humoral and cellular immune responses were assessed in each animal against the antigen with which they were immunized and in each animal both type of responses was detected as shown in Table 11.

All chimpanzees became HCV-RNA positive (determined with Monitor HCV, Roche, Basel, Switzerland) on day 7 post challenge and a first ALT and gammaGT peak was measured between days 35 and 63. This evidences that all chimps developed acute hepatitis. Remarkably, both EI immunized animals resolved their infection while the E2deltaHVRI and the control animal did not. This is evidenced by the fact that the EI immunized animals lost HCV-RNA (determined with Monitor HCV, Roche, Basel, Switzerland) at day 98 (Yoran) and 133 (Marti) and remained negative so far until day 273 with monthly testing. All the other animals stayed RNA-positive during the entire follow up period of 273 days so far with ALT and gammaGT values not returning to normal as for the EI immunized chimpanzees but gradually increasing.

**Table 11:** antibody titers were determined by ELISA two weeks after the 6<sup>th</sup> immunization. A serial dilution of the sample was compared to an in house standard (this in house standard defined as having 1000 mU/ml of EI or anti-E2deltaHVR I antibody is a mixture of three sera from HCV chronic carriers selected based on a high anti-envelope titer). The stimulation index, which reflects the cellular immune response, was obtained by culturing PBMC, drawn from the animals two weeks after the third immunization, in the presence or absence of envelope antigen and determining the amount of tritiated thymidine

incorporated in these cells during a pulse of 18 hours after 5 days of culture. The stimulation index is the ratio of thymidine incorporated in the cells cultured with envelope antigen versus the ones cultured without antigen. A stimulation index of >3 is considered a positive signal.

5

Chimpanzee	Anti-E1 response		Anti-E2deltaHVRI response	
	Antibody titer	Stimulation index	Antibody titer	Stimulation Index
Yoran	14110	10.9		
Marti	5630	14.2		
Joost			3210	8.5
Karlien			1770	11.2

Three weeks after the last of 6 immunizations all animals including the control were challenged with 100 CID (chimpanzee infectious doses) of a genotype 1b inoculum (J4.91, kindly provided by Dr. J. Bukh, NIH, Bethesda, Maryland). The amino acid sequence divergence between the vaccine proteins and the J4.91 isolate (of which the sequence information is available under accession number BAA01583) is 7% (9 out of 135 amino acids) for E1s and 11% (32 out of 304 amino acids) for E2delta HVRI; Consequently this challenge is considered heterologous and reflects a real life challenge.

15 In conclusion we have shown that E1-immunization changes the natural history of HCV infection by preventing evolution to a chronic infection, which is the major health problem related with HCV.

**Example 16: Similar E1 responses which allowed clearing of infection in chimpanzee can be induced in man**

In order to obtain a prophylactic effect of E1 immunization in man it is required that similar immune responses can be induced in man compared to chimpanzee. Therefore we vaccinated 20 male human volunteers, in which no anti-E1 responses (humoral or cellular) could be detected, with 3 doses of 20 µg E1s formulated on 0.13% Alhydrogel in 0.5 ml.

25 All immunizations were given intramuscularly with a 3 week interval. As evidenced in

Table 12, 17 out of 20 volunteers indeed mounted a significant humoral and cellular immune response against EI and this without serious adverse events. Only 1 volunteer (subject 021) should be considered as a non-responder since neither humoral nor cellular responses were above the cut-off level after 3 EI immunizations. The observation that the

5 humoral response is lower compared to chimpanzee relates to the fact that only 3 immunizations with 20 µg were given and not 6 with 50 µg.

**Table 12:** antibody titers were determined by ELISA two weeks after the third immunization. A serial dilution of the sample was compared to an in house standards (this in house standard defined as having 1000 mU/ml of E1 or anti-E2deltaHVR I antibody is a mixture of three sera from HCV chronic carriers selected based on a high anti-envelope titer). The stimulation index (cellular immune response) was obtained by culturing PBMC, drawn from the individuals two weeks after the third immunization, in the presence or absence of 1 µg of E1s and determining the amount of tritiated thymidine incorporated in these cells during a pulse of 18 hours after 5 days of culture. The stimulation index is the ratio of thymidine incorporated in the cells cultured with envelope antigen versus the ones cultured without antigen. A stimulation index of >3 is considered a positive signal.

Subject no	Antibody titer	Stimulation index	Subject no	Antibody titer	Stimulation index
002	1370	30.9	014	49	4.6
003	717	13.2	015	228	3.8
004	800	9.1	016	324	4.1
007	680	3.8	017	<20*	6.2
008	1026	3.9	018	<20	6.7
009	325	4.6	019	624	3.1
010	898	7.7	020	84	5.5
011	284	4.1	021	<20	2.1
012	181	3.6	022	226	2.7
013	<20	3.5	023	163	7.6

\*this individual is considered anti-E1 positive after immunization since a significant increase in ELISA signal was seen between the preimmune sample and the sample after three immunization, the titer however is very low and does not allow accurate determination.



**Example 17: Boosting of E1 responses in vaccinated healthy volunteers**

19 out of the 20 human volunteers of example 16 were boosted once more with 20 µg E1s formulated on 0.13% Alhydrogel in 0.5 ml at week 26 (i.e. 20 weeks after the third immunization). Again antibody titers and cellular immune responses were determined 2 weeks after this additional immunization. In all individuals the antibody titer had decreased during the 20 week interval but could easily be boosted by this additional immunization to a level equal or higher of that observed at week 8. On average the antibody titer was double as high after this boost compared to the week 8 titer, and 7 times as high compared to the week 26 titer (Table 13).

Remarkably the T-cell responses were for the majority of individuals still high after the 20 week interval. Taking in account a normalization to the tetanos response, which is present in most individuals as a consequence of previous vaccinations, there is no change in the geometric mean of the stimulation index. After the additional boost, taking in account a normalization to the tetanos response, no change is noted (figure 51). This confirms that a strong T-help response was induced after 3 E1 immunizations and indicates that these immunizations induced already a very good T-help memory which requires, at least for a period of 6 months, no further boosting.

20

25

30

**Table 13:** antibody titers were determined by ELISA two weeks (= week 8) and 20 weeks (= week 26) after the third immunization and finally also 2 weeks after the boost (= Week 28). A serial dilution of the sample was compared to an in house standards (this in house standard defined as having 1000 mU/ml of E1 antibody, is a mixture of three sera from HCV chronic carriers selected based on a high anti-envelope titer). For accurate comparison the determination of the titer at week 8 was repeated within the same assay as for the week 26 and 28 samples, which explains the differences with table 12 of example 16.

10

Subject no	Antibody titer		
	Week 8	Week 26	Week 28
002	1471	443	3119
003	963	95	2355
004	1006	409	2043
007	630	65	541
008	926	81	819
009	704	77	269
010	1296	657	3773
011	253	65	368
012	254	148	760
013	36	<20	166
014	53	40	123
016	159	45	231
017	109	39	568
018	43	23	50
019	425	157	1894
020	73	33	113
021	25	<20	26
022	280	150	357
024	177	81	184
average	467	138	936

Legend to figure 51: The stimulation index (cellular immune response) was obtained by culturing PBMC ( $10^5$  cells), drawn from the individuals before immunization (week 0), two weeks after the third immunization (week 8), before the booster immunization (week 26) and two weeks after the booster immunization (week 28), in the presence or absence of 3  $\mu$ g of recombinant E1s or 2  $\mu$ g tetanos toxoid and determining the amount of tritiated thymidine incorporated in these cells during a pulse of 18 hours after 5 days of culture. The stimulation index is the ratio of thymidine incorporated in the cells cultured with envelope antigen versus the ones cultured without antigen. Samples of week 0 and 8 were determined in a first assay (A), while the samples of week 26 and 28 were determined in a second assay (B) in which the samples of week 0 were reanalyzed. Results are expressed as the geometric mean stimulation index of all 20 (A, experiment) or 19 (B, experiment) volunteers.

In addition the Th1 cytokine interferon-gamma and Th2 cytokine interleukin-5 were measured in the supernatants of the PBMC cultures of samples taken at week 26 and 28 and restimulated with E1. As can be judged from figure 52 the predominant cytokine secreted by the E1 stimulated PBMC is interferon-gamma. It is highly surprising to see that a strong Th1 biased response is observed with an alum adjuvanted E1, since alum is known to be a Th2 inducer. Once more the results confirm that a good T-cell memory response is induced, as prior to the final boost (week 26) already a very strong response is observed. The interferon-gamma secretion was found to be specific as in an additional experiment we saw no difference in interferon-gamma secretion between E1 stimulated cell cultures and non-stimulated cell cultures of these volunteers using samples drawn at week 0.

Legend to figure 52: PBMC ( $10^5$  cells), drawn from the individuals before the booster immunization (week 26) and two weeks after the booster immunization (week 28), were cultured in the presence of 3  $\mu$ g of recombinant E1s (E1) or 2  $\mu$ g of tetanos toxoid (TT) or no antigen (BI). Cytokines were measured in the supernatant taken after 24 hours (interleukin-5) or after 120 hours (interferon-gamma) by means of ELISA. The stimulation index is the ratio of cytokine measured in the supernatants of cells cultured with envelope antigen versus the ones cultured without antigen. Results are expressed as the geometric mean of pg cytokine/ml secreted of all 19 volunteers. Samples with a cytokine amount below detection limit were assigned the value of the detection limit. Similarly samples with

extremely high concentrations of cytokine out of the linear range of the assay were assigned the value of the limit of the linear range of the assay.

**Example 18: Fine mapping of cellular response against E1 in vaccinated healthy volunteers.**

In order to map the E1 specific responses a series of 20-mer peptides was synthesized, using standard Fmoc chemistry, with 8 amino acids overlap and covering the entire sequence of E1s. All peptides were C-terminally amidated and N-terminally acetylated, with the exception of IGP 1626 which has a free amino-terminus.

10

IGP1626: YEVRNVSGIYHVTNDCSNSS (amino acid 192-211)(SEQ ID NO:112)

IGP1627: TNDCSNSSIVYEAADMIMHT (amino acid 204-223)(SEQ ID NO:113)

IGP1628: AADMIMHTPGCVPCVRENNNS (amino acid 216-235)(SEQ ID NO:114)

IGP1629: PCVRENNSSRCWVALTPTLA (amino acid 228-247)(SEQ ID NO:115)

15 IGP1630: VALTPTLAARNASVPTTTIR (amino acid 240-259)(SEQ ID NO:116)

IGP1631: SVPTTTIRRHVDLLVGAAAF (amino acid 252-271)(SEQ ID NO:117)

IGP1632: LLVGAAAFCSAMYVGDLGCS (amino acid 264-283)(SEQ ID NO:118)

IGP1633: YVGDLGCSVFLVSQLFTISP (amino acid 276-295)(SEQ ID NO:119)

IGP1634: SQLFTISPRRHETVQDCNCS (amino acid 288-307)(SEQ ID NO:120)

20 IGP1635: TVQDCNCSIYPGHITGHRMA (amino acid 300-319)(SEQ ID NO:121)

IGP1636: HITGHRMAWDMMMNWSPTTA (amino acid 312-331)(SEQ ID NO:122)

PBMC from 14 different healthy donors not vaccinated with E1s or 10 donors vaccinated with E1s were cultured in the presence of 25 µg/ml (non vaccinated persons) or 10 µg/ml (vaccinated persons, samples taken after the third or booster injection) of each peptide separately. As can be judged from figure 53 the peptides IGP 1627, 1629, 1630, 1631, 1633, 1635 and 1635 all induced significantly higher responses in vaccinated persons compared to non-vaccinated persons. Using a stimulation index of 3 as cut-off the peptides IGP 1627, 1629, 1631 and 1635 were the most frequently recognized (i.e. recognized by at least half of the vaccinated persons tested).

30

This experiment proves that the T-cell responses induced by E1s derived from mammalian cell culture are specific against E1 since these responses can not only be recalled by the same E1s derived from mammalian cell culture but also by synthetic peptides. In addition

this experiment delineates the most immunogenic T-cell domains in E1 are located between amino acids 204-223, 228-271, 276-295, 300-331 and more particularly even between amino acids 204-223, 228-247, 252-271 and 300-319.

- 5 Legend to figure 53: The stimulation index (cellular immune response) was obtained by culturing PBMC ( $3 \times 10^5$  cells), in the presence or absence of peptides and determining the amount of tritiated thymidine incorporated in these cells during a pulse after 5-6 days of culture. The stimulation index is the ratio of thymidine incorporated in the cells cultured with peptide versus the ones cultured without peptide. Results are expressed as individual  
10 values for vaccinated persons (top panel) or non vaccinated or controls (lower panel).

The present invention also provides therefor, the following E1 peptides, proteins, compositions and kits containing the same, nucleic acid sequences coding for these peptides and proteins containing the same, and methods of their manufacture and use, as are  
15 generally described herein for other E1 and related peptides of the present invention.

- |    |          |  |
|----|----------|--|
|    | IGP 1626 | spanning positions 192-211 of the E1 region (SEQ ID NO:112), |
|    | IGP 1627 | spanning positions 204-223 of the E1 region (SEQ ID NO:113), |
|    | IGP 1628 | spanning positions 216-235 of the E1 region (SEQ ID NO:114), |
|    | IGP 1629 | spanning positions 228-247 of the E1 region (SEQ ID NO:115), |
| 20 | IGP 1630 | spanning positions 240-259 of the E1 region (SEQ ID NO:116), |
|    | IGP 1631 | spanning positions 252-271 of the E1 region (SEQ ID NO:117), |
|    | IGP 1632 | spanning positions 264-283 of the E1 region (SEQ ID NO:118), |
|    | IGP 1633 | spanning positions 276-295 of the E1 region (SEQ ID NO:119), |
|    | IGP 1634 | spanning positions 288-307 of the E1 region (SEQ ID NO:120), |
| 25 | IGP 1635 | spanning positions 300-319 of the E1 region (SEQ ID NO:121), |
|    | IGP 1636 | spanning positions 312-331 of the E1 region (SEQ ID NO:122). |

**Example 19: Exemplification of therapeutic utility of E1-vaccination in man**

- The HCV E1s protein (amino acids 192-326 (SEQ ID NO: 123):  
30 YEVRNVSGMYHVTNDCSNSSIVYEAADMIMHTPGCVPCVRENNSSRCWVALTPT  
LAARNASVPTTTIRRHVDLLVGAAAFCSAMYVGDLGSLVLSQLFTISPRRHETV  
QDCNCSIYPGHITGHRMAWDMMNW)) was expressed in Vero cells using  
recombinant vaccinia virus HCV11B. This vaccinia virus is essentially identical to

vvHCV11 A (described in, for example, PCT/EP95/03031 and U.S. Patent No. 6,510,134, the entire contents of each of which are incorporated herein by reference) but has been passed from RK13 to Vero cells. The protein was purified (by means of lentil chromatography, reduction-alkylation and size exclusion chromatography) essentially as described in Example 9 of PCT/E99/04342 (the entire contents of each of which are incorporated herein by reference), making use of iodoacetamide as alkylating agent for the cysteines. After purification the 3% Empigen-BB (N-Dodecyl-N,N-dimethylglycine) was exchanged to 3% betain by size exclusion chromatography as described in Example 1 of PCT/E99/04342 this process allows to recover E1s as a particle. Finally the material was desalted to PBS containing 0.5% betain and an E1s concentration of 400 µg/mL. This E1 was mixed with an equal volume of Alhydrogel 1.3% (Superfos, Denmark) and finally further diluted with 8 volumes of 0.9% NaCl to yield alum-adjuvanted E1 at a concentration of 40 µg E1/mL and 0.13% of Alhydrogel.

In order to demonstrate a therapeutic effect of E1 immunization in man, an immune response was induced in chronically infected patients. This immune response was both quantitatively and qualitatively different from the base line response against E1 which is present in such patients.

Twenty-six chronically genotype 1 infected HCV patients were vaccinated with 5 doses of 20 µg E1s formulated on 0.13% Alhydrogel in 0.5 mL. All immunizations were given intramuscularly. Immunizations were performed at week 0, 4, 8, 12 and 24. In addition, 9 patients received an identical number of placebo injections consisting of alum only.

As evidenced in the following Table 14, a T-cell response is normally absent in chronic HCV carriers (only 4%, 1 out of 26, have a detectable T-cell reactivity). Upon immunization this increases to about 70% (18 out of 26) already after 4 immunizations. This observation demonstrates that the immune response has been qualitatively changed from a T-cell non-responsiveness to a clear response. Furthermore this immune response can be sustained, by boosting with larger intervals since after the fifth immunization 3 months later, the cellular immune response is still of the same level or slightly increased. No significant changes were observed in the placebo group.

**Table 14:** The stimulation index (SI; cellular immune response) was obtained by culturing PBMC, drawn from the individuals four weeks (W16) after the fourth immunization and two weeks (W26) after the fifth immunization in the presence or absence of 3 µg of E1s and determining the amount of tritiated thymidine incorporated in these cells during a pulse

of 18 hours after 5 days of culture. The stimulation index is the ratio of thymidine incorporated in the cells cultured with envelope antigen versus the ones cultured without antigen. A stimulation index of  $>3$  is considered a positive signal.

	T-cell response		
	W0	W16	W26
SI $>3$	1/26	18/26	17/25
Mean SI	1.5	13.7	19.3

- 5 In addition the antibody titers were determined by ELISA prior to the first immunization (this was done for each patient on three samples taken at different time points before immunization) and after the fifth immunization (this was done for each patient on two samples: W26 and W28). A serial dilution of the sample was compared to an in-house standard (this in-house standard is, as previously described, defined as having 1000  
10 mU/mL of E1 antibody is a mixture of three sera from HCV chronic carriers selected based on a high anti-envelope titer). From this analysis it was concluded that, on average, the titer doubled from 331 to 715 mU/mL. This result demonstrates that the humoral arm of the immune response has at least been quantitatively altered.

- 15 An E1 vaccine formulated on alum significantly changes both the qualitative and quantitative immune response against E1. Such an E1 based vaccine and/or any of the vaccines described herein may be additionally useful if administered in conjunction with antiviral therapy such but not limited to interferon and alternatively with its combination with ribavirin (i.e., prior to, after or with a composition of the present invention).

- 20 The following as well as all previously and subsequently identified text, references, patents, publication, etc., are hereby incorporated herein in their entirety by reference.

**Example 20: Effects of therapeutic E1-vaccination in chronically infected patients.**

- In a first course of this study, 26 patients received 5 doses of 20  $\mu$ g E1s as described in Example 19. In a second course, 25 of said 26 patients received a further 6 intramuscular  
25 doses of 20  $\mu$ g E1s formulated on 0.13% Alhydrogel in 0.5 mL (as in Example 19). The immunizations of this second series were administered with 3-week intervals and the first injection was given at week 50. Four weeks after the last injection of the second course, a liver biopsy was performed in 24 out of the 25 patients who completed the two courses of therapeutic E1s vaccination. A liver biopsy was also performed in all patients prior to the

onset of the therapeutic vaccine treatment (i.e. prior to course 1 vaccination scheme). The mean time elapsed between pre-treatment and post-treatment biopsies was 17 months. Liver histology slides thus obtained before and after therapeutic vaccination were scored centrally in a blind way by two expert pathologists for fibrosis and inflammation, this  
5 according to the Ishak and Metavir scoring systems (Ishak et al. 1995 which is a modification of the scoring system of Knodell et al. 1981; Bedossa and Poynard 1996), as well as for anti-E2 immunostain using the IGH222 murine anti-E2 HVRI monoclonal antibody as described in International Patent Application WO99/50301. Perisinusoidal fibrosis was assessed based on staining of collagen with Sirius Red (on liver histology  
10 slides).

The Ishak scores range from 0 to 18 for grading of inflammation and from 0 to 6 for staging of fibrosis/cirrhosis. The sum of the Ishak inflammation and fibrosis scores comes closest to the Histological Activity Index (HAI; Knodell et al. 1981) which has been  
15 widely used. The change in Ishak fibrosis score in the 24 patients who followed the two courses of therapeutic E1s vaccination is  $-0.04$  (with a 95% confidence interval of  $-0.60$  to  $+0.68$ ) and with a change from baseline- to end-value from  $2.54$  (baseline) to  $2.50$  (end). An overview of the different assessed necro-inflammatory intensities (Ishak scoring) is given in Table 15.

20 The Ishak equivalent of the HAI scores for the treated patients showed a mean absolute change from baseline of  $-0.17$  (with a 95% confidence interval of  $-1.36$  to  $1.03$ ) and with a change from baseline- to end-value from  $8.88$  (baseline) to  $8.71$  (end). Furthermore, nine (9) out of the 24 patients (38%) improved 2 points or more on the sum of Ishak inflammation- and Ishak fibrosis- scores whereas ten (10) out of the 24 patients remained  
25 stable (no change or a change of  $+1$  or  $-1$ ) and five (5) patients evolved to a worse condition (worsening of 2 points or more).

The Metavir scores range from 0 to 3 for grading of inflammation and from 0 to 4 for staging of fibrosis/cirrhosis. The overall progression rate of the Metavir score in an  
30 untreated patient is estimated to be 0.133 per year (Poynard et al. 1997).

For the patients of the current study, the average progression over the 17-month period based on the linear extension of the baseline score and the estimated duration of the infection (which is reported for 19 out of the 24 patients who followed the two courses of



therapeutic E1s vaccination) would be 0.20. This correlates well with the published overall rate of progression which would be 0.19 for the 17 month period.

The change in Metavir score observed for the treated patients is, however, 0.00 (with a 95% confidence interval of -0.43 to +0.43) and with a mean baseline- an end-score of 1.67.

- 5 An overview of Metavir scores at baseline and at the end of the second course treatment is given in Table 16. Unblinded comparison of pre- and post treatment slides revealed that 10 patients had diminished perisinusoidal fibrosis based on Sirius Red staining.

- 10 The anti-E2 immunostain scores range from 0 to 4, with 0 = no E2 antigen detectable, 1 = occasional cell with detectable E2 antigen, 2= clusters of E2 antigen positive cells but less than 25% of cells positive, 3 = clusters of E2 antigen positive cells and 25-50% of cells positive, 4 = clusters of E2 antigen positive cells and more than 50% of cells positive. The anti-E2 immunostain scores for the treated patients showed a mean absolute change from baseline of -0.75 (with a 95% confidence interval of -1.64 to 0.14) and with a change from  
15 baseline- to end-value from 2.54 (baseline) to 1.79 (end). The reduction seen for anti-E2 immoreactivity in the liver using blind scoring was confirmed by unblinded comparison of paired slides per patient. Eleven patients showed negativation or a marked reduction in HCV E2 immunostaining, while only three patients had a stronger immunostain after treatment. Of these 11 patients, 3 also showed a reduction in steatosis.

20

- Serum HCV RNA levels were determined using the Amplicor HCV Monitor kit (Roche, Basel, Switzerland). The serum HCV RNA levels remained unchanged or did not change more than one log from baseline except in 1 patient. This concerned a 37-year old treatment-naïve female patient, infected with a genotype 1a virus and with flu-like  
25 symptoms at baseline. HCV-RNA dropped 3 logs from week 8 (after two injections with E1) to reach levels below 3000 IU/ml at weeks 16 and 20, in parallel ALT dropped from a single peak value of 400 U/ml at week 8 to normal ALT values at weeks 16, 20 and 24. This was accompanied by disappearance of the flu-like symptoms. After this period both ALT and HCV-RNA increased again, and symptoms re-appeared. The sequence of the  
30 virus present during the emerging viral load increase was compared with the baseline sequence. No evidence for the presence of escape mutant virus was found.

Serum ALT (alanine aminotransferase activity) levels in the treated patients decreased on average 22% (with a 95% confidence interval of -15% to -30%) as compared to the baseline levels. A positive correlation between % change from baseline in ALT levels and absolute change from baseline in fibrosis scores was noticed: a Spearman rank of  $p=0.0007$  for Ishak fibrosis score and of  $p=0.002$  for Metavir fibrosis score were obtained as correlations between the change from baseline in serum ALT and the absolute change from baseline Ishak and Metavir fibrosis scores, respectively.

A further outcome of the current study is the predictive character of the immune response (in terms of increase in serum anti-E1s antibody levels) to treatment for improvement in histological scores (fibrosis and overall) as well as serum ALT levels.

Antibody titers were determined by ELISA. A serial dilution of a serum sample was compared to an in house standard (this in house standard defined as having 1000 mU/mL of E1s antibody is a mixture of three sera from HCV chronic carriers selected based on a high anti-envelope titer). The detection limit for this assay is 5 mU/mL.

A Spearman rank of  $p=0.007$  was obtained as correlation between the change from baseline in serum anti-E1 antibody level and the absolute change from baseline in total Ishak (inflammation plus fibrosis) score. Pearson  $p$ -values of  $p=0.06$  and of  $p=0.009$  were obtained as correlation between the change from baseline in serum anti-E1 antibody level and the absolute ( $p=0.06$ ) and relative ( $p=0.009$ ) change from baseline in Ishak fibrosis score, respectively. A Pearson  $p$ -value of  $p=0.01$  was obtained as correlation between the change from baseline in serum anti-E1 antibody level and the relative change from baseline in ALT value.

These associations remain after correction for baseline liver histology scores, ALT, serum HCV-RNA viral load, age and sex, and IFN exposure and thus as add to the likelihood of a real positive effect induced by the therapeutic E1s vaccination treatment.

The seven patients with the highest increase in serum antibody levels in response to the therapeutic vaccine candidate (i.e., patients with an increase in anti-E1 antibody level of at least 700 mU/mL) had on average a significant decrease of -0.9 points (with a 95% confidence interval of -0.2 to -1.5) on Metavir liver fibrosis score. In this subgroup, serum ALT on average decreased 37% (with a 95% confidence interval of -25 to -49%). Table 17 provides an overview of the correlation between immune response (in terms of anti-E1s

antibody levels induced by treatment) and overall change (relative to baseline) in Ishak scores. Figure 54 illustrates the correlation between changes in Ishak fibrosis score, changes in ALT-levels and changes in anti-E1s antibody levels. Figure 55 illustrates the correlation between changes in Metavir fibrosis score, changes in ALT-levels and changes in anti-E1s antibody levels. Figure 56 illustrates Ishak fibrosis score versus ALT levels both at the start of the first treatment course (panel A) and at the end of the second treatment course (panel B). Figure 57 illustrates patient's age versus Ishak fibrosis score both at the start of the first treatment course (panel A) and at the end of the second treatment course (panel B). Both Figures 56 and 57 further indicate the seven patients with the highest increase in anti-E1s antibody levels induced by the therapeutic vaccination treatment.

A positive correlation between serum anti-E1s antibody levels and T cell proliferation index values at the end of the second treatment course was observed, with a Pearson p-value of  $p=0.009$ .

Concludingly, the increase in antibody levels to the E1s vaccination significantly predicted improvement in liver fibrosis (Metavir and Ishak scores), sum of inflammation and fibrosis scores (Ishak), and in ALT levels, even after correcting for any other baseline prognostic variables. The increase in anti-E1 antibodies after vaccination was also significantly correlated with an increase in T cell proliferation index to E1.

This study is clearly supportive for a E1-based therapeutic vaccination strategy to have the potential to halt disease progression towards liver cirrhosis.

**Table 15.** Ishak grading of necro-inflammatory intensities for periportal hepatitis, confluent necrosis, focal inflammation, portal inflammation and the overall total inflammation grading. Scores are indicated as the change from baseline (mean and 95 % confidence interval) and the mean baseline- to end-values.

Score for	Change of score from baseline mean (95% confidence interval)	Change of score from baseline to end (mean)
periportal hepatitis	-0.21 (-0.62 to 0.20)	1.42 to 1.21
confluent necrosis	0.38 (-0.17 to 0.92)	0.00 to 0.38
focal inflammation	-0.17 (-0.59 to 0.26)	2.63 to 2.46
portal inflammation	-0.13 (-0.51 to 0.26)	2.29 to 2.17
total inflammation grading	-0.13 (-1.18 to 0.93)	6.33 to 6.21

- Table 16.** Overview of frequencies (given as number of patients) of changes of a given baseline Metavir score (given in top row; Baseline 0 to 4) to a given Metavir score at the end of the second course of therapeutic EIs vaccination (given in left row; EOT 0 to 4). For instance, the "5" marked with a "\*" (i.e., "5\*") means that 5 patients had a baseline Metavir score of 1 and a Metavir score of 0 at the end of the second course treatment (EOT = end of treatment).
- 5

	Baseline 0	Baseline 1	Baseline 2	Baseline 3	Baseline 4	Total
EOT 0	0	5*	1	0	0	6
EOT 1	1	3	2	0	0	6
EOT 2	0	5	1	0	0	6
EOT 3	0	0	0	1	1	2
EOT4	0	0	1	2	1	4
<b>Total</b>	<b>1</b>	<b>13</b>	<b>5</b>	<b>3</b>	<b>2</b>	<b>24</b>

**Table 17.** Correlation between serum anti-E1 antibody levels induced by therapeutic E1s vaccination and change in overall Ishak scores. Given are the number of patients corresponding to the possible criteria as outlined in the Table.

Overall change of Ishak score (change from baseline)	< -1	-1, 0, or +1	> +1
High anti-E1s antibody response	5	2	0
Low anti-E1s antibody response	4	8	5

**Example 21: Effects of therapeutic E1-vaccination regimen.**

In a first course of this study, 26 patients received 5 doses of 20 µg E1s ("E1s patients") and 9 patients received an identical number of placebo injections ("placebo patients") consisting of alum only, as described in Example 19. In a second course, 25 of said 26 patients that received E1s during the first course and 9 of said 9 patients that received a placebo were all immunized with a further 6 intramuscular doses of 20 µg E1s formulated on 0.13% Alhydrogel in 0.5 mL (as in Example 19), thus giving rise to two groups of patients: the "E1s/E1s patients" and "placebo/E1s patients". The immunizations of this second series, i.e. course 2, were administered with 3-week intervals (compared to 4-week intervals during the first series, i.e. course 1) and the first injection was given at week 50.

The median values for anti-E1 antibodies after 4 E1s injections were 195 mU/mL for 4-week intervals (as deduced from the E1s patients) and 274 mU/mL for 3-week intervals (as deduced from the placebo/E1s patients), thus indicating that a 3-week interval regimen induces a stronger humoral immune response in a shorter period of time.

In placebo/E1s patients E1s-specific T cell proliferation increased from 0/9 patients at baseline to 9/9 patients after 6 injections with 3-week intervals at week 69.

If the humoral and cellular immune response to E1s in healthy male volunteers (see Examples 16 and 17) are compared with those in the patient group, a somewhat slower increase in anti-E1 antibody levels and T cell response is observed. The booster effect seen in healthy volunteers was almost absent in patients (median titer of 165 mU/mL after the fifth immunization administered to the E1s/E1s patients as a booster 12 weeks after the 4<sup>th</sup> immunization), thus suggesting a somewhat impaired build up of immunological memory to E1s in patients. Repeated intramuscular injections at 3-week intervals seem to overcome this impairment (median titer of 530 mU/mL after 6 consecutive immunizations with a 3-week

interval in the placebo/E1s patients), thus indicating another advantage of an E1s vaccination regimen with 3-week intervals and/or administering more than 4 doses in the first series of immunizations.

## REFERENCES

- Alric L, Duffaut M, Selves J, Sandre K, Mularczyck M, Izopet J, Desmorat H, Bureau C, Chaouche N, Dalbergue B, Vinel JP (2001) *J Hepatol* **35**(2):272-278.
- 5 Asseline U, Delarue M, Lancelot G, Toulme F, Thuong N (1984) *Proc. Natl. Acad. Sci. USA* **81**:3297-301.
- Bailey, J. and Cole, R. (1959) *J. Biol. Chem.* **234**, 1733-1739.
- Ballou, L., Hitzeman, R., Lewis, M. & Ballou, C. (1991) *PNAS* **88**, 3209-3212.
- Bedossa P, Poynard T. (1996) *Hepatology* **24**(2):289-293.
- 10 Benesch, R., Benesch, R.E., Gutcho, M. & Lanfer, L. (1956) *Science* **123**, 981.
- Cavins, J. & Friedman. (1970) *Anal. Biochem.* **35**, 489.
- Cleland, W. (1964) *Biochemistry* **3**, 480
- Creighton, E. (1988) *BioEssays* **8**, 57
- Darbre, A., John Wiley & Sons Ltd. (1987) *Practical Protein Chemistry - A Handbook*.
- 15 Darbre, A., John Wiley & Sons Ltd. (1987) *Practical Proteinchemistry* p.69-79.
- Doms et al, (1993), *Virology* **193**, 545-562.
- Ellman, G. (1959) *Arch. Biochem. Biophys.* **82**, 70.
- Falck-Ytter Y, Kale H, Mullen KD, Sarbah SA, Sorescu L, McCullough AJ (2002) *Ann Intern Med* **136**(4):288-292.
- 20 Falkner, F. & Moss, B. (1988) *J. Virol.* **62**, 1849-1854.
- Friedman, M. & Krull. (1969) *Biochem. Biophys. Res. Commun.* **37**, 630.
- Gallagher J. (1988) *J. Cell Biol.* **107**, 2059-2073.
- Glazer, A., Delange, R., Sigman, D. (1975) North Holland publishing company, Elsevier, *Biomedical. Part : Modification of protein* (p. 116).
- 25 Graham, F. & van der Eb, A. (1973) *Virology* **52**, 456-467.
- Grakoui et al. (1993) *Journal of Virology* **67**:1385-1395.
- Grassetti, D. & Murray, J. (1969) *Analyt. Chim. Acta.* **46**, 139.
- Grassetti, D. & Murray, J. (1967) *Arch. Biochem Biophys.* **119**, 41.
- Helenius, *Mol. Biol. Cell* (1994), **5**: 253-265.
- 30 Hijikata, M., Kato, N., Ootsuyama, Y., Nakagawa, M. & Shimotohno, K. (1991) *Proc. Natl. Acad. Sci. U.S.A.* **88**(13):5547-51.
- Hochuli, E., Bannwarth, W., Döbeli, H., Gentz, R., Stüber, D. (1988) *Biochemistry* **88**, 8976.

- Hsu, H., Donets, M., Greenberg, H. & Feinstone, S. (1993) *Hepatology* 17:763-771.
- Inoue, Y., Suzuki, R., Matsuura, Y., Harada, S., Chiba, J., Watanabe, Y., Saito, I. & Miyamura, T. (1992) *J. Gen. Virol.* 73:2151-2154.
- Ishak K, Baptista A, Bianchi L, Callea F, De Groote J, Gudat F, Denk H, Desmet V, Korb  
5 G, MacSween RN, et al. (1995) *J Hepatol* 22(6):696-699.
- Janknecht, R., de Martynoff, G. *et al.*, (1991) *Proc. Natl. Acad. Sci. USA* 88, 8972-8976.
- Kayman (1991) *J. Virology* 65, 5323-5332.
- Kato, N., Oostuyama, Y., Tanaka, T., Nakagawa, M., MuraIso, K., Ohkoshi, S., Hijikata, M., Shimitohno, K. (1992) *Virus Res.* 22:107-123.
- 10 Kniskern, P., Hagopian, A., Burke, P., Schultz, L., Montgomery, D., Hurni, W., Yu Ip, C., Schulman, C., Maigetter, R., Wampler, D., Kubek, D., Sitrin, R., West, D., Ellis, R., Miller, W. (1994) *Vaccine* 12:1021-1025.
- Kohara, M., Tsukiyama-Kohara, K., Maki, N., Asano, K., Yoshizawa, K., Miki, K., Tanaka, S., Hattori, N., Matsuura, Y., Saito, I., Miyamura, T. & Nomoto, A. (1992) *J. Gen.*  
15 *Virol.* 73:2313-2318.
- Knodell RG, Ishak KG, Black WC, Chen TS, Craig R, Kaplowitz N, Kiernan TW, Wollman J. (1981) *Hepatology* 1(5):431-435.
- Lauer GM, Walker BD (2001) *N Engl J Med* 345(1):41-52.
- Lefkowitz JH (1997) *Viral Hepatitis Reviews* 3(3):155-166.
- 20 Mackett, M., Smith, G. & Moss, B. (1985) In: 'DNA cloning: a practical approach' (Ed. Glover, D.) IRL Press, Oxford.
- Mackett, M., & Smith, G. (1986) *J. Gen. Virol.* 67, 2067-2082.
- Mackett, M., Smith, G. & Moss, B. (1984) *J. Virol.* 49, 857-864.
- Mackett, M., Smith, G. & Moss, B. (1984) *Proc. Natl. Acad. Sci. USA* 79, 7415-7419.
- 25 Matsukura M, Shinozuka K, Zon G, Mitsuya H, Reitz M, Cohen J, Broder S (1987) *Proc. Natl. Acad. Sci. USA* 84:7706-10.
- Means, G. (1971) Holden Day, Inc.
- Means, G. & Feeney, R. (1971) Holden Day p.105 & p. 217.
- Miller P, Yano J, Yano E, Carroll C, Jayaram K, Ts'o P (1979) *Biochemistry* 18:5134-43.
- 30 Mita, E., Hayashi, N., Ueda, K., Kasahara, A., Fusamoto, H., Takamizawa, A., Matsubara, K., Okayama, H. & Kamada T. (1992) *Biochem. Biophys. Res. Comm.* 183:925-930.
- Moore, S. (1963) *J. Biol. Chem.* 238, 235-237.
- Nielsen P, Egholm M, Berg R, Buchardt O (1991) *Science* 254:1497-500.



- Nielsen P, Egholm M, Berg R, Buchardt O (1993) *Nucleic-Acids-Res.* **21**:197-200.
- Okamoto, H., Okada, S., Sugiyama, Y., Yotsumoto, S., Tanaka, T., Yoshizawa, H., Tsuda, F., Miyakawa, Y. & Mayumi, M. (1990) *Jpn. J. Exp. Med.* **60**:167-177.
- Panicali & Paoletti (1982) *Proc. Natl. Acad. Sci. USA* **79**, 4927-4931.
- 5 Piccini, A., Perkus, M. & Paoletti, E. (1987) *Meth. Enzymol.* **153**, 545-563.
- Poynard T, Bedossa P, Opolon P. (1997) *Lancet* 349(9055):825-832.
- Poynard T, McHutchison J, Manns M, Trepo C, Lindsay K, Goodman Z, Ling MH, Albrecht J (2002) *Gastroenterology* **122**(5):1303-1313.
- Reichard O, Glaumann H, Fryden A, Norkrans G, Wejstal R, Weiland O (1999) *J Hepatol*
- 10 **30**(5):783-787.
- Rose (1988) *Annu. Rev. Cell Biol.* 1988, **4**: 257-288;
- Ruegg, V. and Rudinger, J. (1977) *Methods Enzymol.* **47**, 111-116.
- Schvarcz R, Glaumann H, Reichard O, Weiland O (1999) *J Viral Hepat* **6**(3):237-242.
- Shan, S. & Wong (1993) *CRC-press* p. 30-33.
- 15 Shiffman ML (1999) *Viral Hepatitis Reviews* **5**(1):27-43.
- Spaete, R., Alexander, D., Rugroden, M., Choo, Q., Berger, K., Crawford, K., Kuo, C., Leng, S., Lee, C., Ralston, R., et al. (1992) *Virology* **188**(2):819-30.
- Skehel, J., (1984) *Proc. Natl. Acad. Sci. USA* **81**, 1179-1783.
- Stunnenberg, H., Lange, H., Philipson, L., Miltenburg, R. & van der Vliet, R. (1988) *Nucl.*
- 20 *Acids Res.* **16**, 2431-2444.
- Stuyver, L., Van Arnhem, W., Wyseur, A., DeLeys, R. & Maertens, G. (1993a) *Biochem. Biophys. Res. Commun.* **192**, 635-641.
- Stuyver, L., Rossau, R., Wyseur, A., Duhamel, M., Vanderborght, B., Van Heuverswyn, H., & Maertens, G. (1993b) *J. Gen. Virol.* **74**, 1093-1102.
- 25 Stuyver, L., Van Arnhem, W., Wyseur, A., Hernandez, F., Delaporte, E., Maertens, G. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 10134-10138.
- Takimoto M, Ohkoshi S, Ichida T, Takeda Y, Nomoto M, Asakura H, Naito A, Mori S, Hata K, Igarashi K, Hara H, Ohta H, Soga K, Watanabe T, Kamimura T (2002) *Dig Dis Sci.* **47**(1):170-176.
- 30 Weil, L. & Seibler, S. (1961) *Arch. Biochem. Biophys.* **95**, 470.
- Yokosuka, O., Ito, Y., Imazeki, F., Ohto, M. & Omata, M. (1992) *Biochem. Biophys. Res. Commun.* **189**:565-571.

WO 96/04385 (PCT/EP95/03031) - Purified Hepatitis C Virus Envelope Proteins for

Diagnostic and Therapeutic Use.

All references cited herein are incorporated in their entirety by reference.

**We claim:**

- 5 1. A therapeutic HCV vaccine composition for reducing liver disease in a chronic HCV-infected mammal.
2. A therapeutic HCV vaccine composition for reducing liver fibrosis progression in a chronic HCV-infected mammal.
- 10 3. A therapeutic HCV vaccine composition for reducing liver fibrosis in a chronic HCV-infected mammal.
4. A therapeutic HCV vaccine composition for reducing liver disease by at least 2 points according to the overall Ishak score in a chronic HCV-infected mammal.
- 15 5. A therapeutic HCV vaccine composition for reducing liver disease by at least 1 point according to the Ishak fibrosis score in a chronic HCV-infected mammal.
6. A therapeutic HCV vaccine composition for reducing serum ALT levels in a chronic HCV-infected mammal.
- 20 7. A therapeutic HCV vaccine composition for reducing steatosis in a chronic HCV-infected mammal.
- 25 8. A therapeutic HCV vaccine composition for reducing anti-E2 immunoreactivity in the liver of a chronic HCV-infected mammal.
9. The therapeutic HCV vaccine composition according to any of claims 1 to 8 which comprises a HCV antigen and, optionally, a pharmaceutically acceptable adjuvant.
- 30 10. The therapeutic HCV vaccine composition according to claim 9 wherein said HCV antigen is an E1 or E2 antigen, or an immunogenic part of an E1 or E2 antigen.

11. The therapeutic HCV vaccine composition according to claim 9 wherein said pharmaceutically acceptable adjuvant is alum.
12. A method to predict changes in liver disease in a chronic HCV-infected mammal comprising
- 5       (i)     determining the level of serum anti-E1 antibody level prior to therapeutic vaccination with a HCV vaccine composition comprising an E1 antigen;
- (ii)    determining the level of serum anti-E1 antibody level after therapeutic vaccination with a HCV vaccine composition comprising an E1 antigen;
- 10       (iii)   inferring the difference in level of serum anti-E1 antibody level determined in (i) and (ii) and, therefrom, predicting the change in liver disease.
13. A therapeutic HCV vaccine composition comprising a therapeutically effective amount of at least one purified HCV single or specific oligomeric recombinant envelope protein selected from the group consisting of an E1 protein, an E2 protein, a part of said
- 15       E1 and E2 proteins, an E1/E2 protein complex formed from purified HCV single or specific oligomeric recombinant E1 or E2 proteins or parts thereof; and optionally a pharmaceutically acceptable adjuvant.
- 20   14. A therapeutic HCV vaccine composition comprising a therapeutically effective amount of a combination of at least two purified HCV single or specific oligomeric recombinant envelope proteins selected from the group consisting of E1 proteins derived from different HCV genotypes or subtypes, E2 proteins derived from different HCV genotypes or subtypes, parts of said E1 and E2 proteins, and E1/E2 protein
- 25       complexes formed from purified HCV single or specific oligomeric recombinant E1 or E2 proteins, or parts thereof, derived from different HCV genotypes or subtypes; and optionally a pharmaceutically acceptable adjuvant.
15. A therapeutic HCV vaccine composition comprising a therapeutically effective amount of at least one of the following E1 and E2 peptides:
- 30       E1-31 spanning amino acids 181 to 200 of the Core/E1-V1 region (SEQ ID NO:56),  
      E1-33 spanning amino acids 193 to 212 of the E1 region (SEQ ID NO:57),  
      E1-35 spanning amino acids 205 to 224 of the E1 V2 region (SEQ ID NO:58),

E1-35A spanning amino acids 208 to 227 of the E1 V2 region (SEQ ID NO:59),  
1bE1 spanning amino acids 192 to 228 of E1 regions V1, C1, and V2 regions (SEQ ID NO:53),  
E1-51 spanning amino acids 301 to 320 of the E1 region (SEQ ID NO:66),  
5 E1-53 spanning amino acids 313 to 332 of the E1 C4 region (SEQ ID NO:67),  
E1-55 spanning amino acids 325 to 344 of the E1 region (SEQ ID NO:68),  
Env 67 or E2-67 spanning amino acid positions 397 to 418 of the E2 region (SEQ ID NO:72),  
Env 69 or E2-69 spanning amino acid positions 409 to 428 of the E2 region (SEQ ID NO:73),  
10 Env 23 or E2-23 spanning positions 583 to 602 of the E2 region (SEQ ID NO:86),  
Env 25 or E2-25 spanning positions 595 to 614 of the E2 region (SEQ ID NO:87),  
Env 27 or E2-27 spanning positions 607 to 626 of the E2 region (SEQ ID NO:88),  
Env 17B or E2-17B spanning positions 547 to 586 of the E2 region (SEQ ID NO:83),  
15 Env 13B or E2-13B spanning positions 523 to 542 of the E2 region (SEQ ID NO:82),  
IGP 1626 spanning positions 192-211 of the E1 region (SEQ ID NO:112),  
IGP 1627 spanning positions 204-223 of the E1 region (SEQ ID NO:113),  
IGP 1628 spanning positions 216-235 of the E1 region (SEQ ID NO:114),  
IGP 1629 spanning positions 228-247 of the E1 region (SEQ ID NO:115),  
20 IGP 1630 spanning positions 240-259 of the E1 region (SEQ ID NO:116),  
IGP 1631 spanning positions 252-271 of the E1 region (SEQ ID NO:117),  
IGP 1632 spanning positions 264-283 of the E1 region (SEQ ID NO:118),  
IGP 1633 spanning positions 276-295 of the E1 region (SEQ ID NO:119),  
IGP 1634 spanning positions 288-307 of the E1 region (SEQ ID NO:120),  
25 IGP 1635 spanning positions 300-319 of the E1 region (SEQ ID NO:121) and  
IGP 1636 spanning positions 312-331 of the E1 region (SEQ ID NO:122);  
and, optionally, a pharmaceutically acceptable adjuvant.

16. A therapeutic HCV composition comprising a therapeutically effective amount of at  
30 least one purified HCV single or specific oligomeric recombinant envelope protein  
selected from the group consisting of an E1 protein, an E2 protein, a part of said E1 and  
E2 protein, and an E1/E2 protein complex formed from purified HCV single or specific  
oligomeric recombinant E1 or E2 proteins or parts thereof; and optionally a

pharmaceutically acceptable adjuvant.

17. A therapeutic HCV composition comprising a therapeutically effective amount of a combination of at least two purified HCV single or specific oligomeric recombinant envelope proteins selected from the group consisting of E1 proteins derived from different HCV genotypes or subtypes, E2 proteins derived from different HCV genotypes or subtypes, parts of said E1 and E2 proteins, and E1/E2 protein complexes formed from purified HCV single or specific oligomeric recombinant E1 or E2 proteins, or parts thereof, derived from different HCV genotypes or subtypes; and optionally a pharmaceutically acceptable adjuvant.
18. A therapeutic HCV composition comprising a therapeutically effective amount of at least one of the following E1 and E2 peptides:
- E1-31 spanning amino acids 181 to 200 of the Core/E1 V1 region (SEQ ID NO:56),
  - E1-33 spanning amino acids 193 to 212 of the E1 region (SEQ ID NO:57),
  - E1-35 spanning amino acids 205 to 224 of the E1 V2 region (SEQ ID NO:58),
  - E1-35A spanning amino acids 208 to 227 of the E1 V2 region (SEQ ID NO:59),
  - 1bE1 spanning amino acids 192 to 228 of E1 regions V1, C1, and V2 regions (SEQ ID NO:53),
  - E1-51 spanning amino acids 301 to 320 of the E1 region (SEQ ID NO:66),
  - E1-53 spanning amino acids 313 to 332 of the E1 C4 region (SEQ ID NO:67),
  - E1-55 spanning amino acids 325 to 344 of the E1 region (SEQ ID NO:68),
  - Env 67 or E2-67 spanning amino acid positions 397 to 418 of the E2 region (SEQ ID NO:72),
  - Env 69 or E2-69 spanning amino acid positions 409 to 428 of the E2 region (SEQ ID NO:73),
  - Env 23 or E2-23 spanning positions 583 to 602 of the E2 region (SEQ ID NO:86),
  - Env 25 or E2-25 spanning positions 595 to 614 of the E2 region (SEQ ID NO:87),
  - Env 27 or E2-27 spanning positions 607 to 626 of the E2 region (SEQ ID NO:88),
  - Env 17B or E2-17B spanning positions 547 to 586 of the E2 region (SEQ ID NO:83),
  - Env 13B or E2-13B spanning positions 523 to 542 of the E2 region (SEQ ID NO:82),
  - IGP 1626 spanning positions 192-211 of the E1 region (SEQ ID NO:112),
  - IGP 1627 spanning positions 204-223 of the E1 region (SEQ ID NO:113),

IGP 1628 spanning positions 216-235 of the E1 region (SEQ ID NO:114),  
IGP 1629 spanning positions 228-247 of the E1 region (SEQ ID NO:115),  
IGP 1630 spanning positions 240-259 of the E1 region (SEQ ID NO:116),  
IGP 1631 spanning positions 252-271 of the E1 region (SEQ ID NO:117),  
5 IGP 1632 spanning positions 264-283 of the E1 region (SEQ ID NO:118),  
IGP 1633 spanning positions 276-295 of the E1 region (SEQ ID NO:119),  
IGP 1634 spanning positions 288-307 of the E1 region (SEQ ID NO:120),  
IGP 1635 spanning positions 300-319 of the E1 region (SEQ ID NO:121) and  
IGP 1636 spanning positions 312-331 of the E1 region (SEQ ID NO:122);  
10 and, optionally, a pharmaceutically acceptable adjuvant.

19. The therapeutic HCV composition according to any of claims 13 to 18 for at least one  
of inducing HCV-specific antibodies, stimulating T-cell activity and stimulating  
cytokine secretion.

15

20. The therapeutic HCV composition according to any of claims 13 to 18 which is  
therapeutically effective in a HCV carrier infected with a HCV genotype different from  
the HCV genotype or HCV genotypes from which said E1, E2, or E1/E2 protein  
complexes are derived.

20

21. A composition comprising at least one of the following E1 and E2 peptides:

IGP 1626 spanning positions 192-211 of the E1 region (SEQ ID NO:112),  
IGP 1627 spanning positions 204-223 of the E1 region (SEQ ID NO:113),  
IGP 1628 spanning positions 216-235 of the E1 region (SEQ ID NO:114),  
25 IGP 1629 spanning positions 228-247 of the E1 region (SEQ ID NO:115),  
IGP 1630 spanning positions 240-259 of the E1 region (SEQ ID NO:116),  
IGP 1631 spanning positions 252-271 of the E1 region (SEQ ID NO:117),  
IGP 1632 spanning positions 264-283 of the E1 region (SEQ ID NO:118),  
IGP 1633 spanning positions 276-295 of the E1 region (SEQ ID NO:119),  
30 IGP 1634 spanning positions 288-307 of the E1 region (SEQ ID NO:120),  
IGP 1635 spanning positions 300-319 of the E1 region (SEQ ID NO:121) and  
IGP 1636 spanning positions 312-331 of the E1 region (SEQ ID NO:122);  
and, optionally, a pharmaceutically acceptable adjuvant.

22. The composition according to any of claims 10 and 13 to 21 wherein the cysteines of said E1 or E2 antigen, of said recombinant HCV envelope proteins or of said E1 and E2 peptides are blocked.
- 5
23. The composition according to any of claims 10, 13, 14, 16, 17, 19, 20 and 22 wherein said E1 or E2 antigen or said recombinant HCV envelope proteins are added as viral-like particles.
- 10
24. The composition according to any of claims 10, 13, 14, 16, 17, 19, 20 and 22 wherein said E1 antigen or said recombinant HCV E1 envelope protein is an E1s protein.
25. The composition according to claim 24 wherein said E1 antigen or said E1s protein is defined by SEQ ID NO:123.
- 15
26. The composition according to any of claims 1-10, 13, 14, 16, 17, 19, 20 and 22 to 25 wherein said HCV antigen or said recombinant HCV envelope proteins are produced by recombinant mammalian cells, by recombinant yeast cells or by a recombinant virus.
- 20
27. The composition according to any of claims 15 and 18 to 25 wherein said peptides are recombinant peptides or synthetic peptides.
28. The composition according to any one of claims 1 to 11 and 13 to 27 for treating a
- 25
- mammal infected with HCV.
29. The composition according to any of claims 1 to 11 and 28 wherein said mammal is a human.
- 30
30. The method according to claim 12 wherein said mammal is a human.



1 / 69

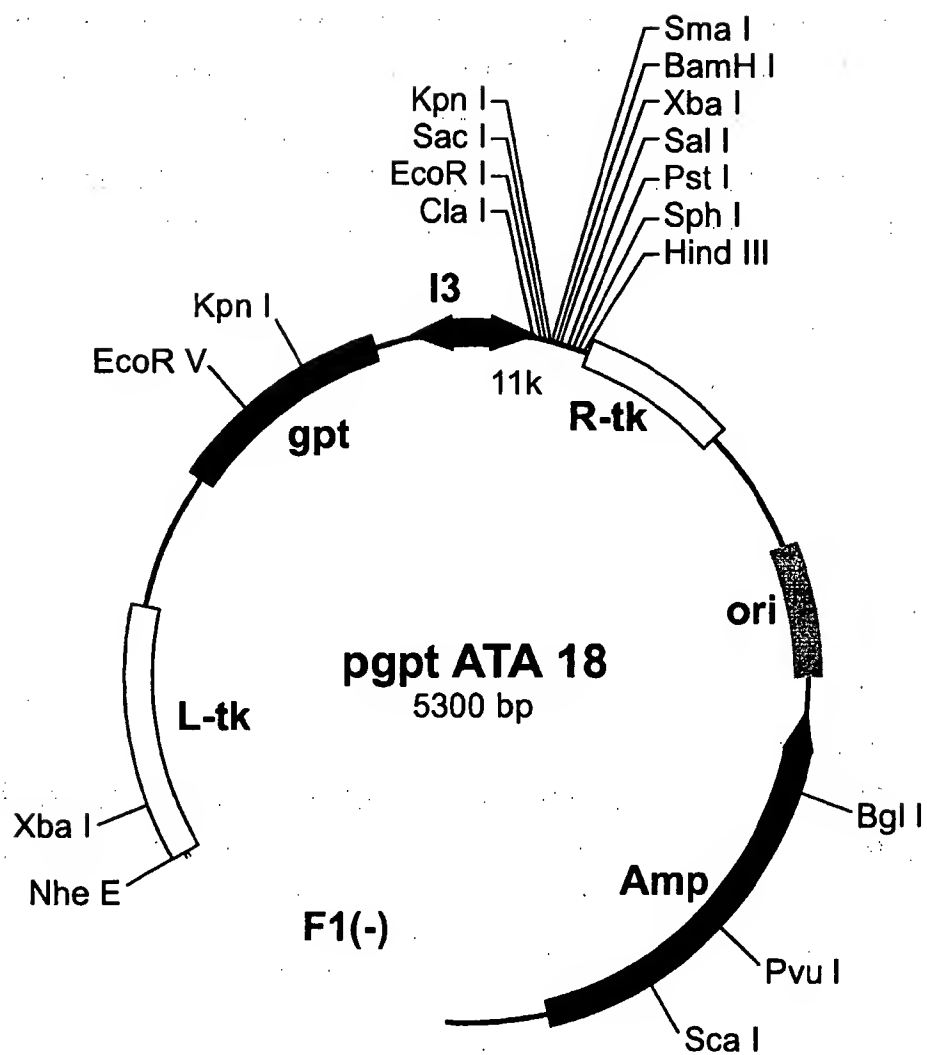


Figure 1

2 / 69

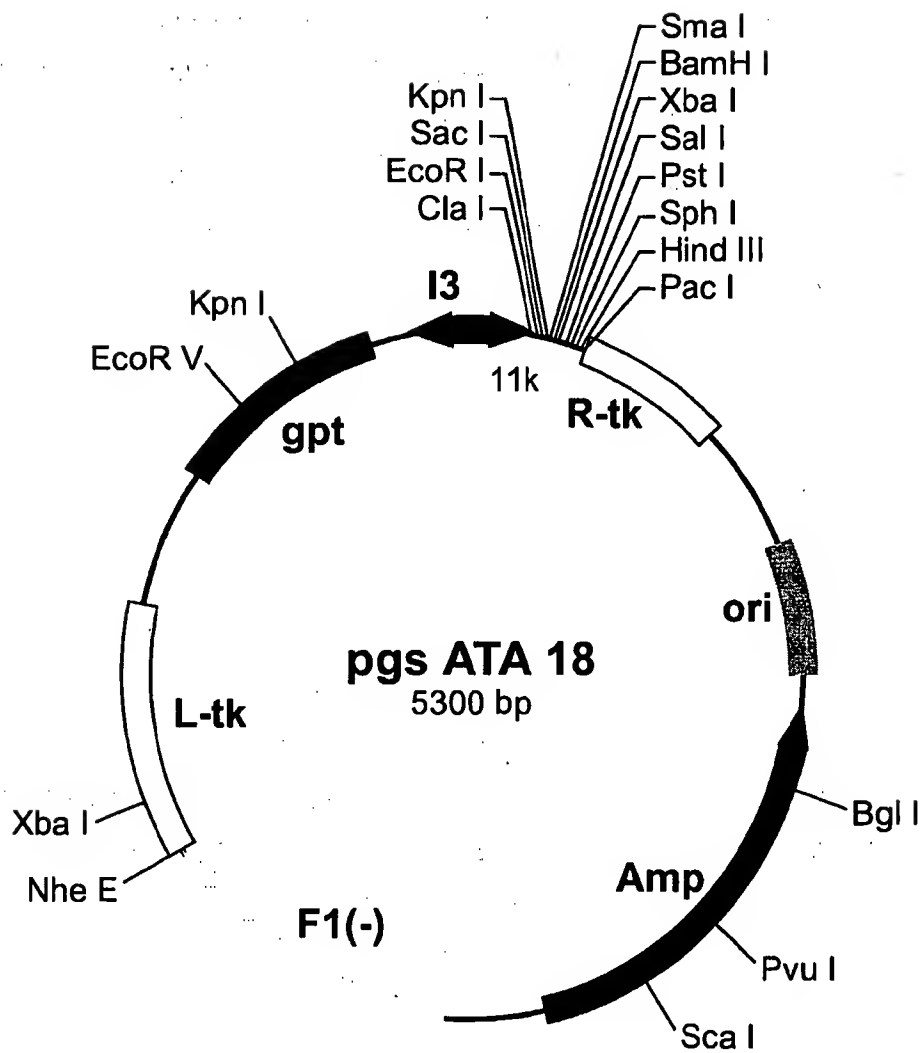


Figure 2

3 / 69

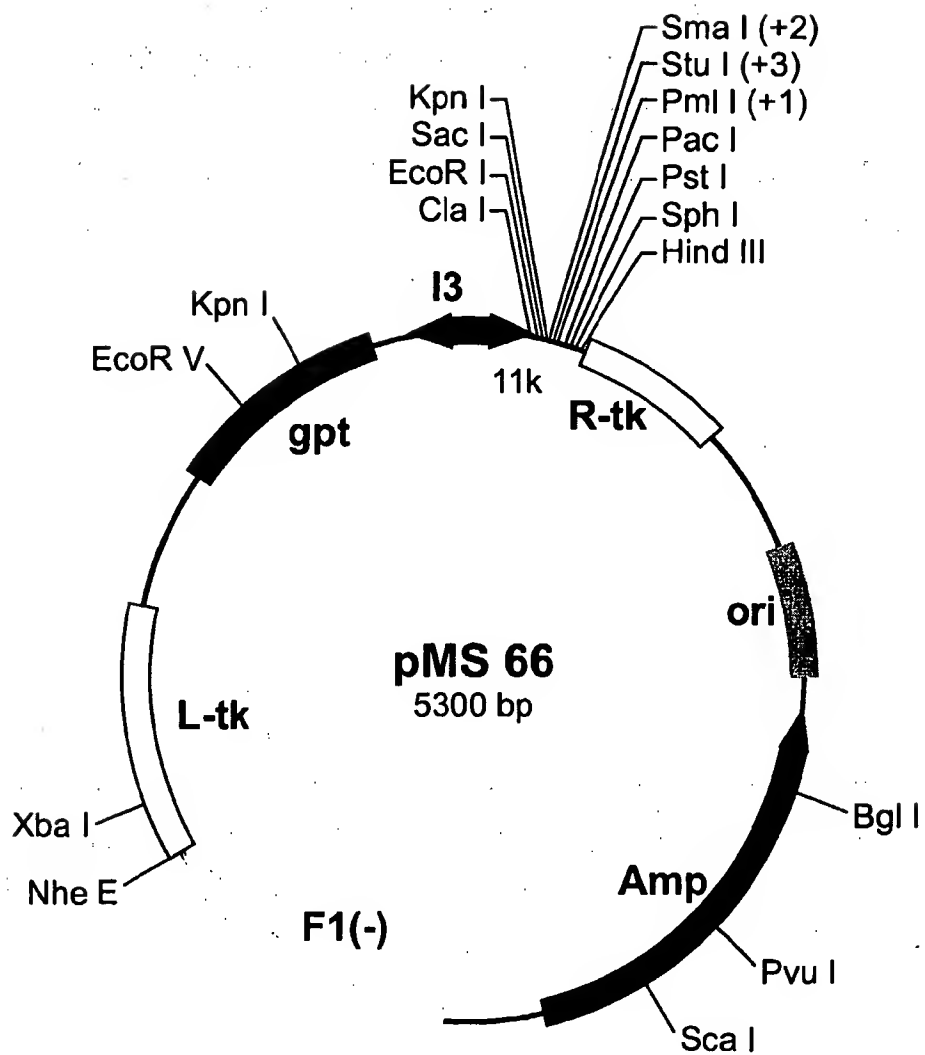


Figure 3

4 / 69

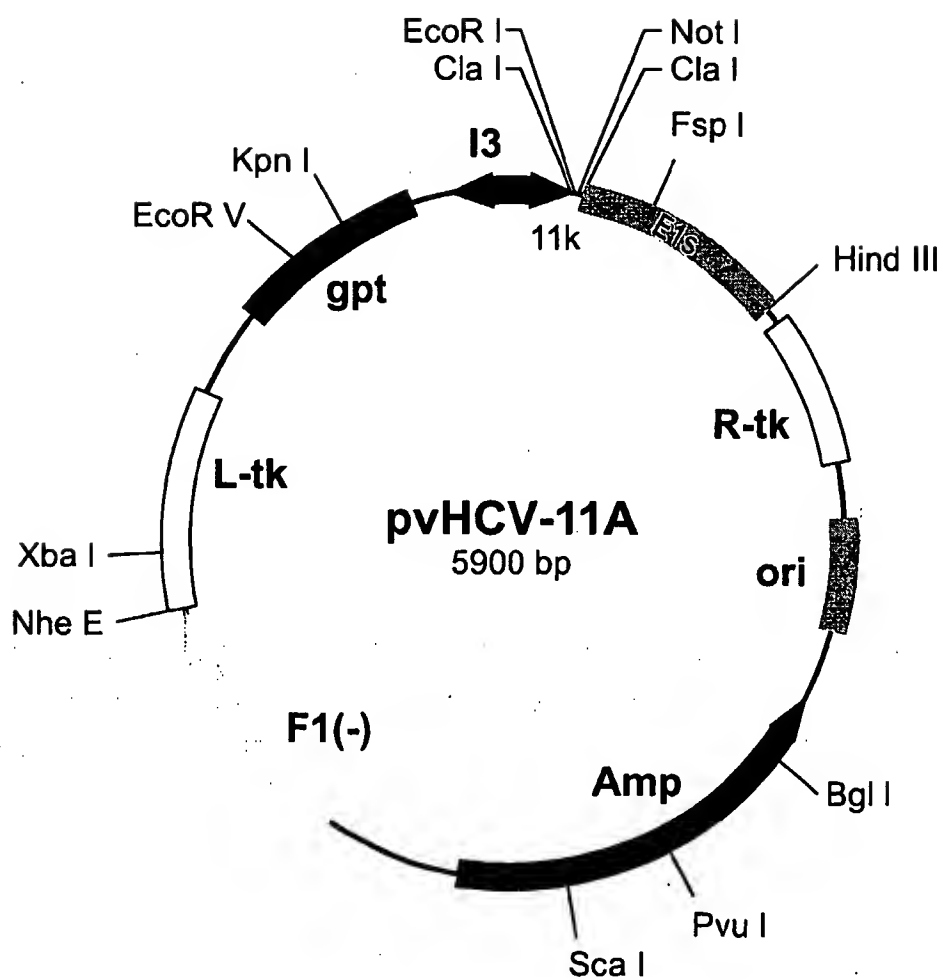


Figure 4

5 / 69

## Anti-E1 levels in NON-responders to IFN treatment

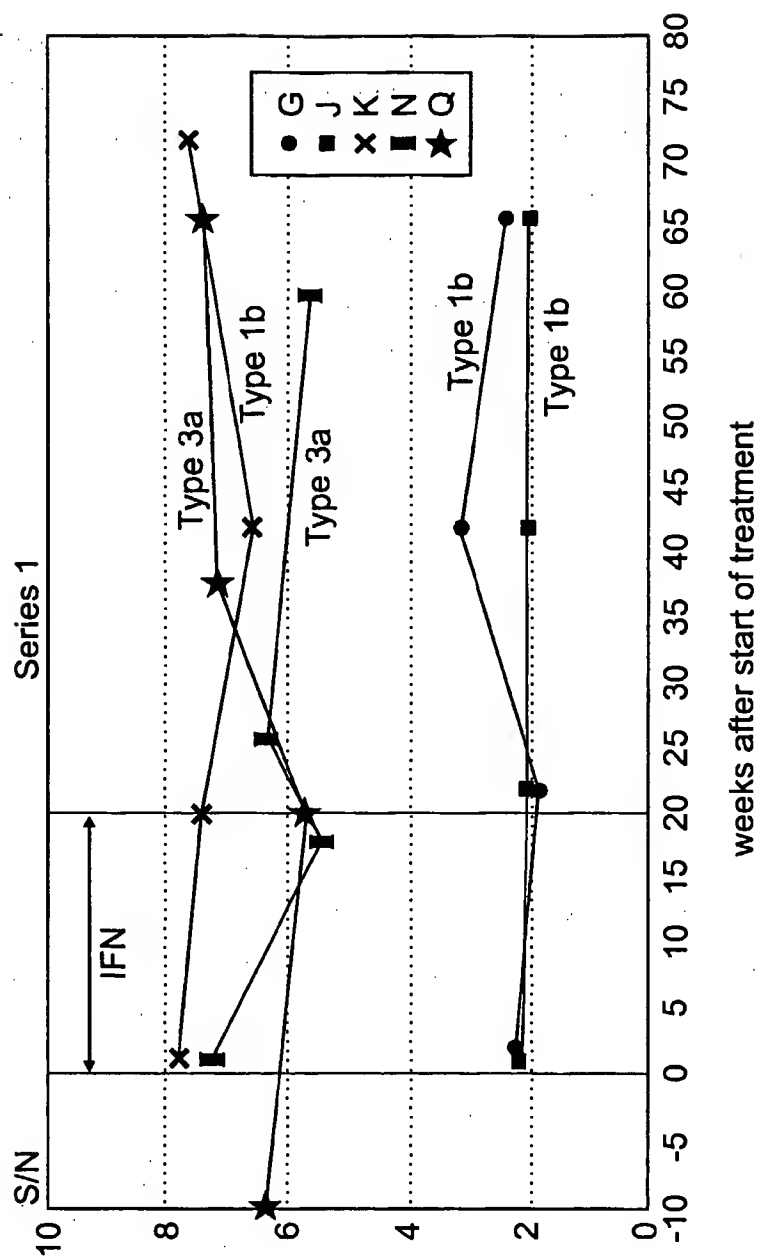


Figure 5

6 / 69

# Anti-E1 levels in RESPONDERS to IFN treatment

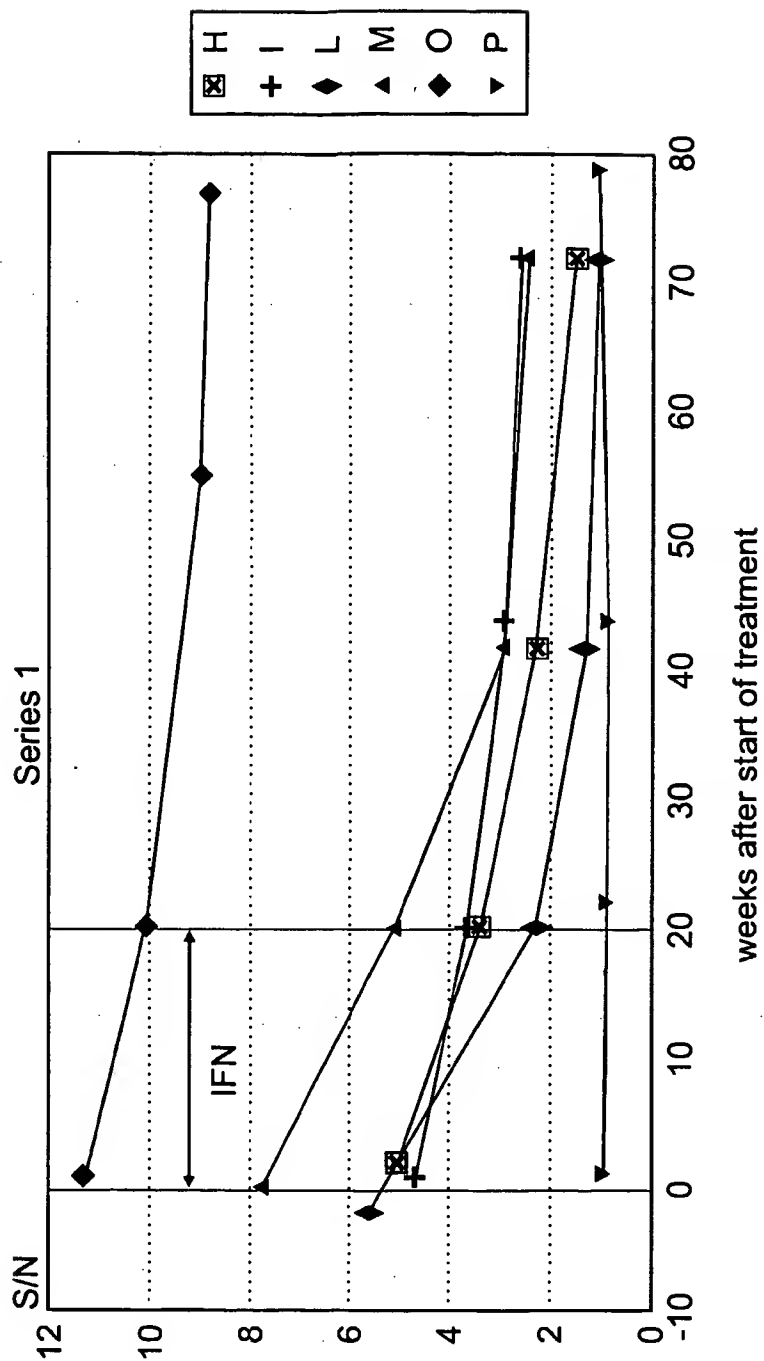


Figure 6

7/ 69

# Anti-E1 levels in patients with COMPLETE response to IFN

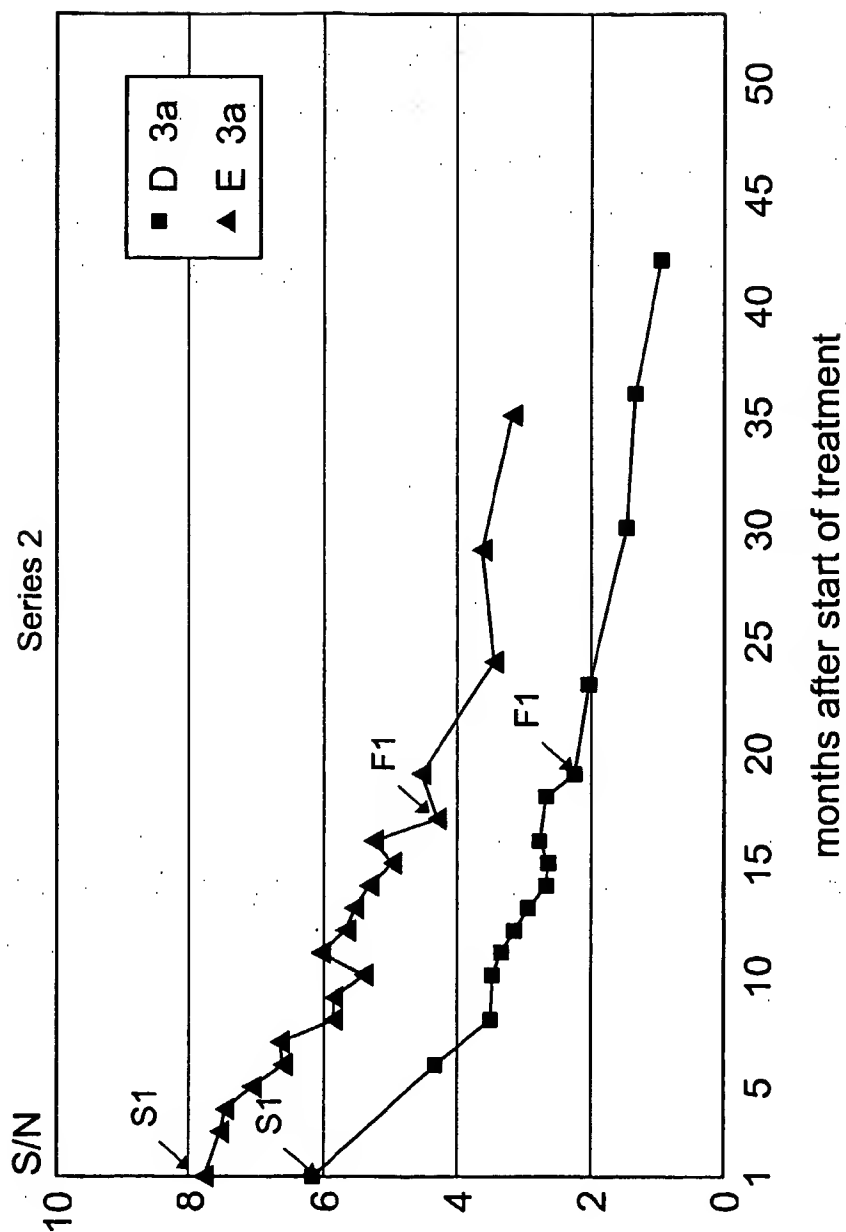


Figure 7

# Anti-E1 levels in INCOMPLETE responders to IFN treatment

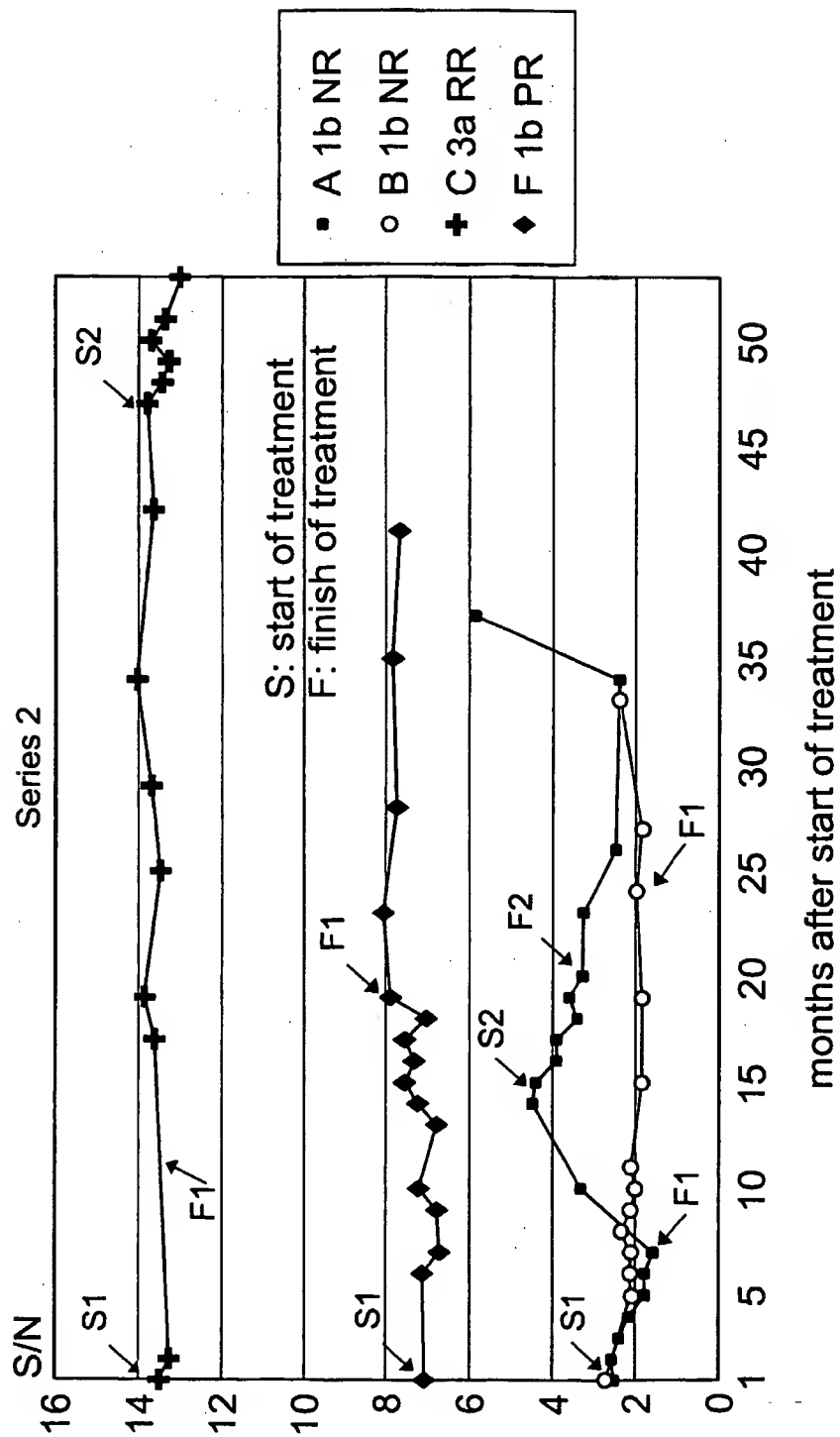


Figure 8



9 / 69

# Anti-E2 levels in RESPONDERS to IFN treatment

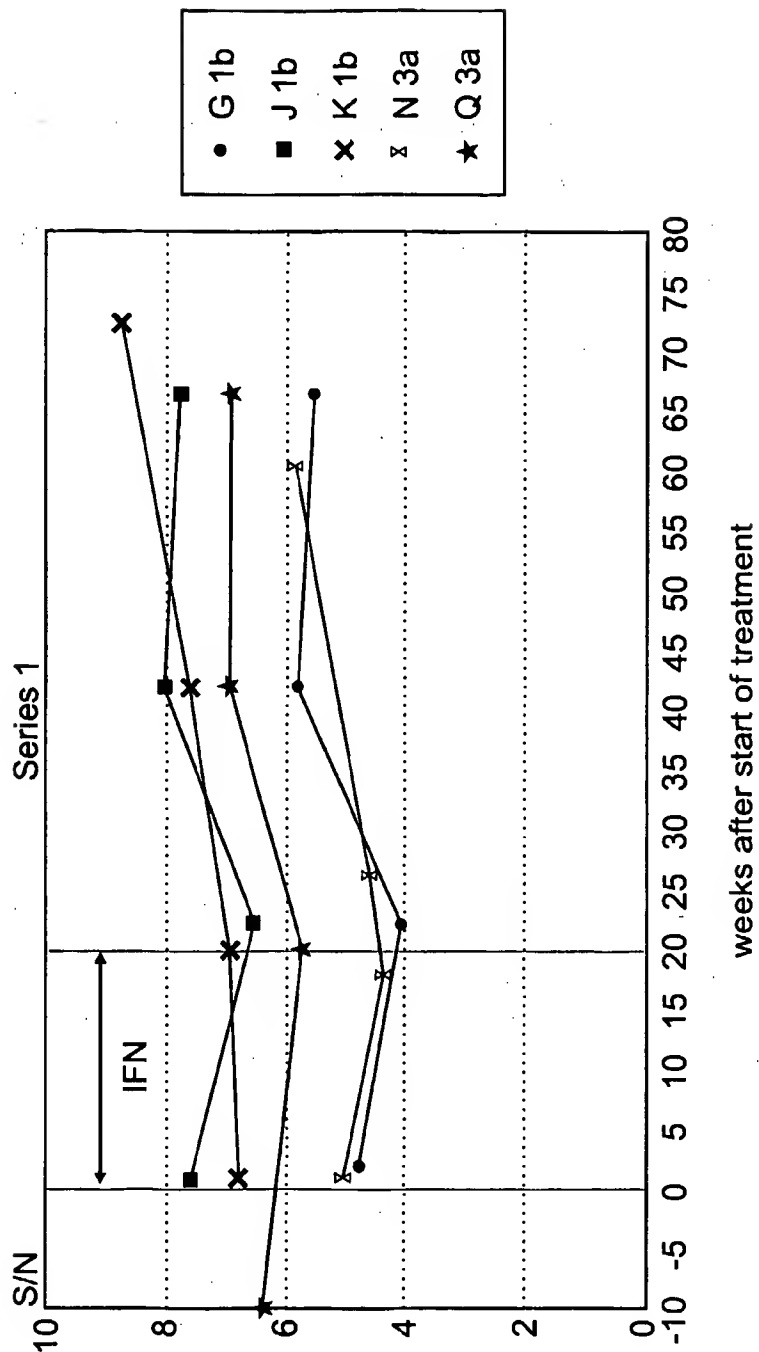


Figure 9

10 / 69

# Anti-E2 levels in RESPONDERS to IFN treatment

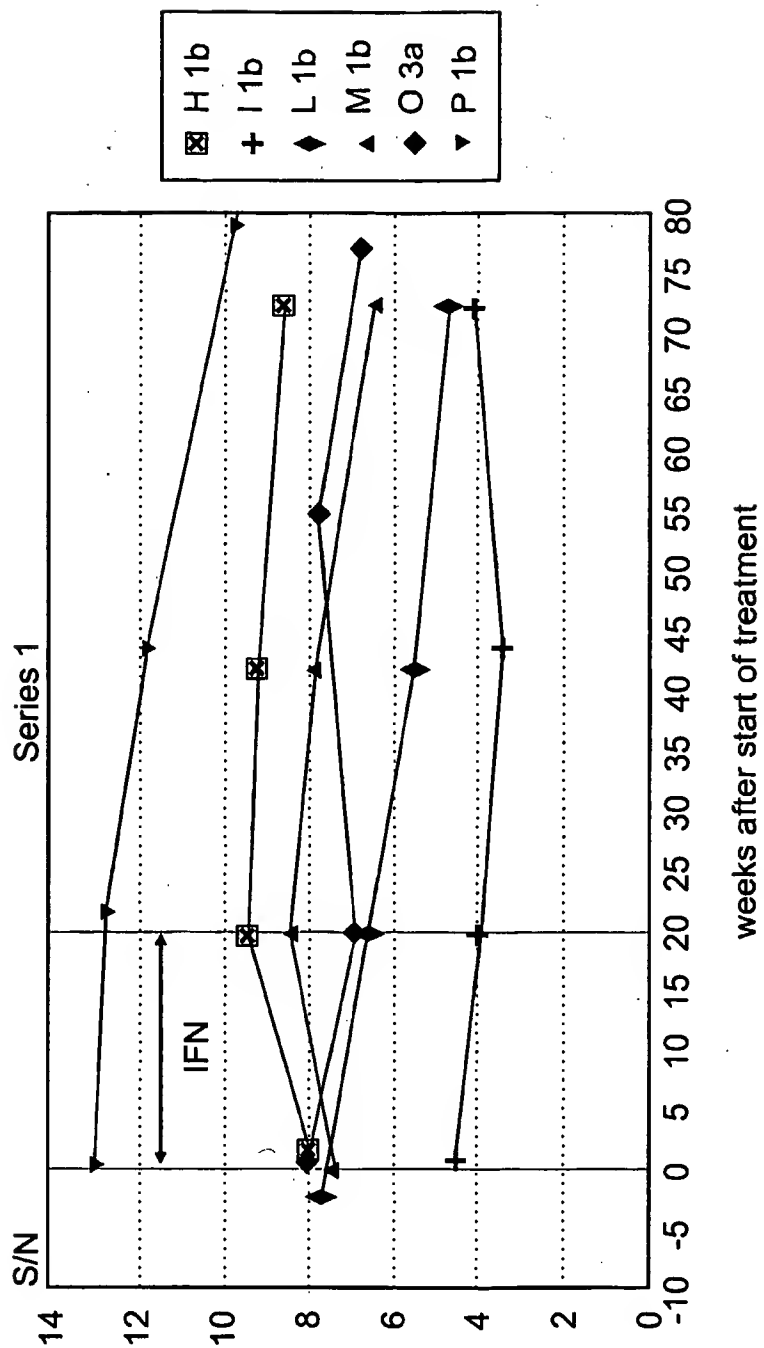


Figure 10

11 / 69

Anti-E2 levels in INCOMPLETE responders to IFN treatment

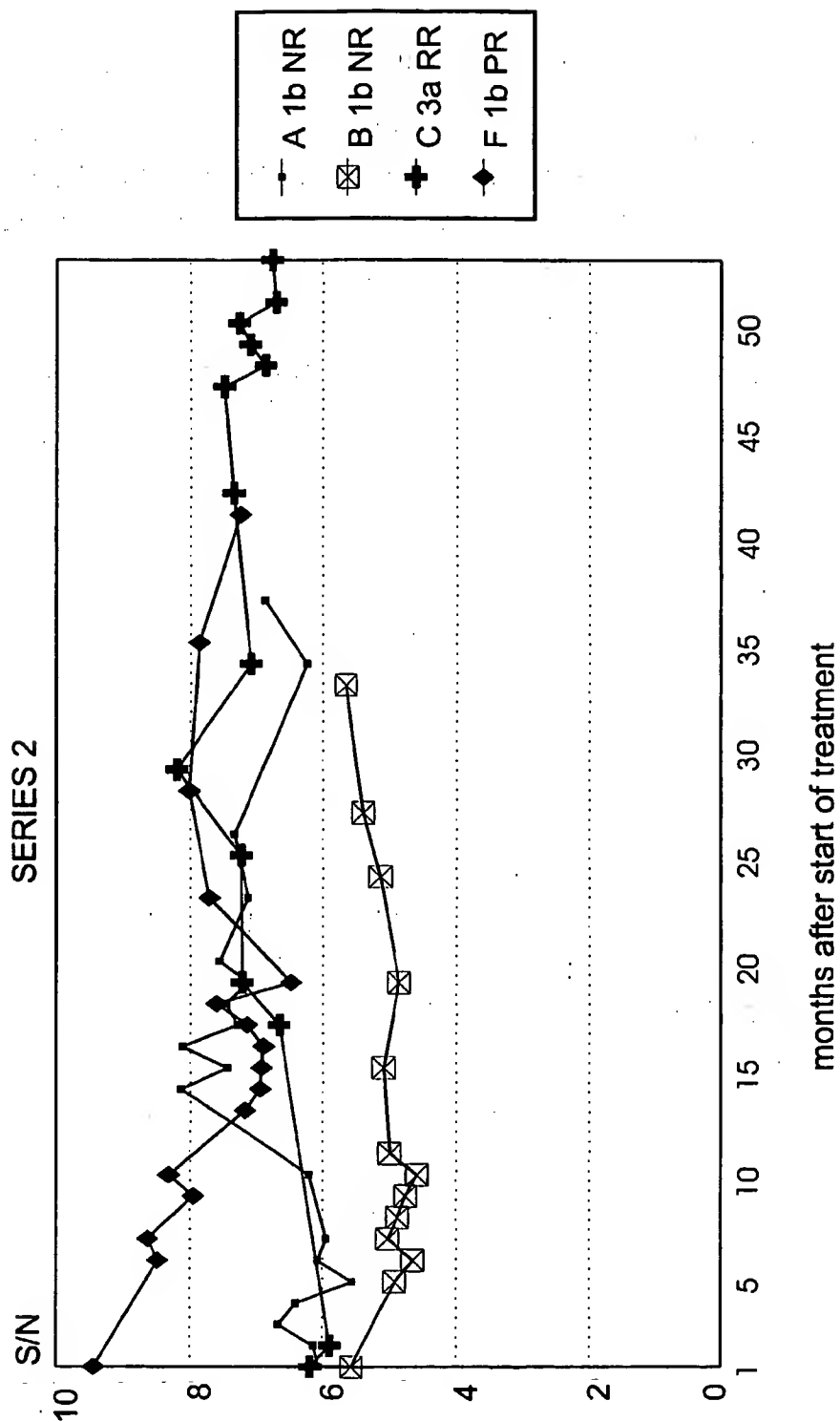


Figure 11

12 / 69

# Anti-E2 levels in COMPLETE responders to IFN treatment

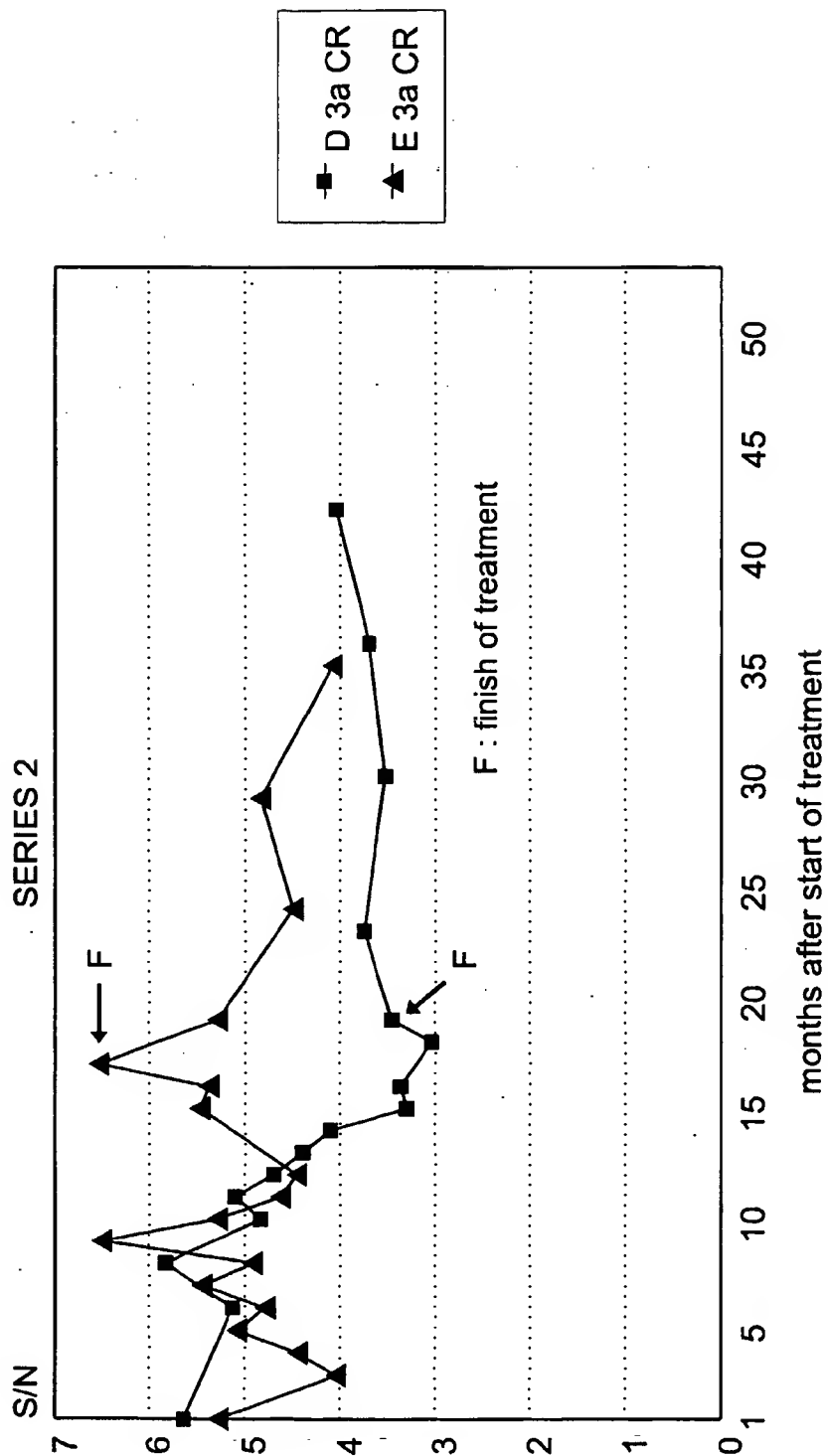


Figure 12

13 / 69

# Human anti-E1 reactivity competed with peptides

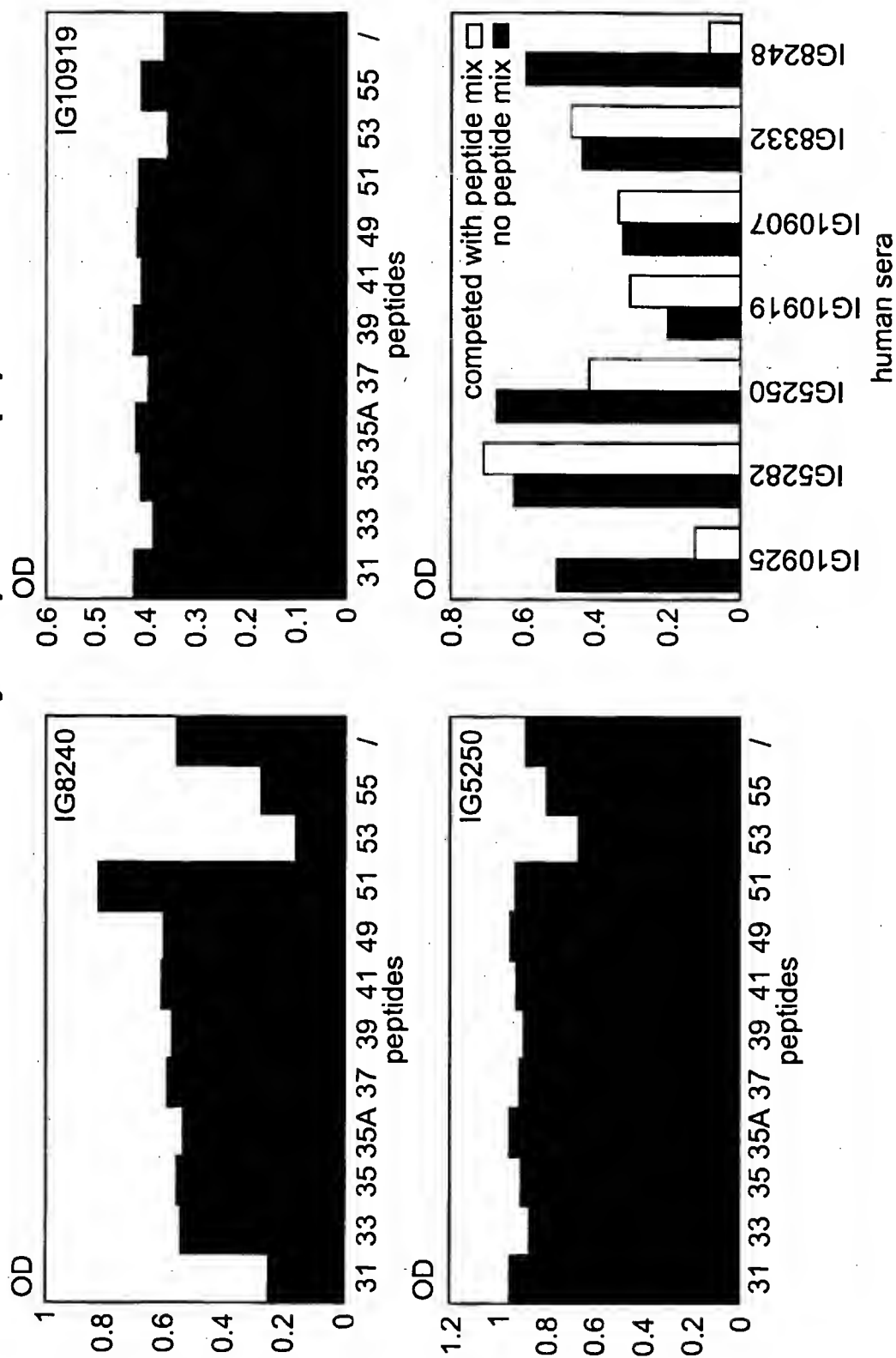


Figure 13

14 / 69

### Competition of reactivity of anti-E1 Mabs with peptides

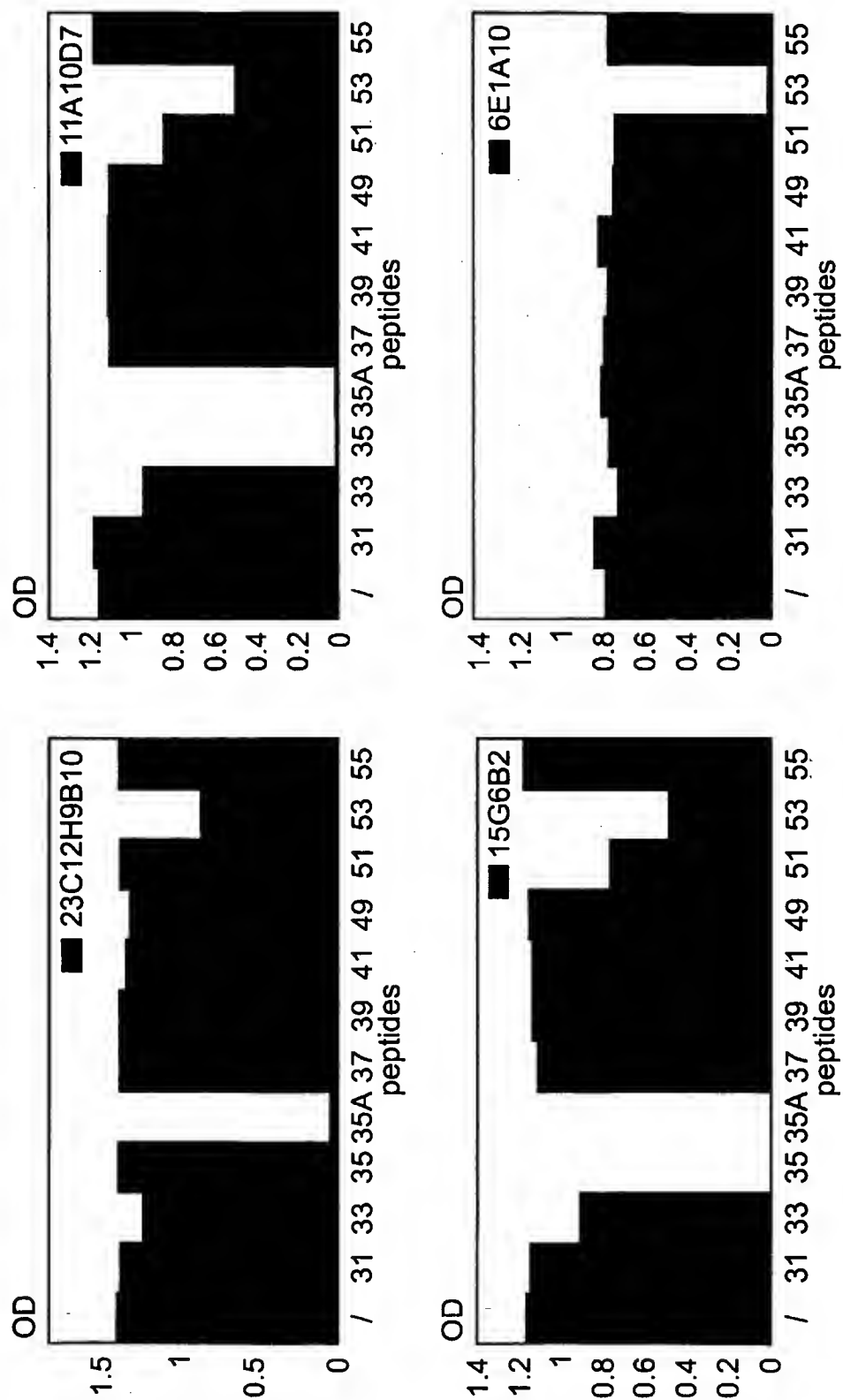


Figure 14

15 / 69

Anti-E1 (epitope 1) levels in NON-RESPONDERS to IFN treatment

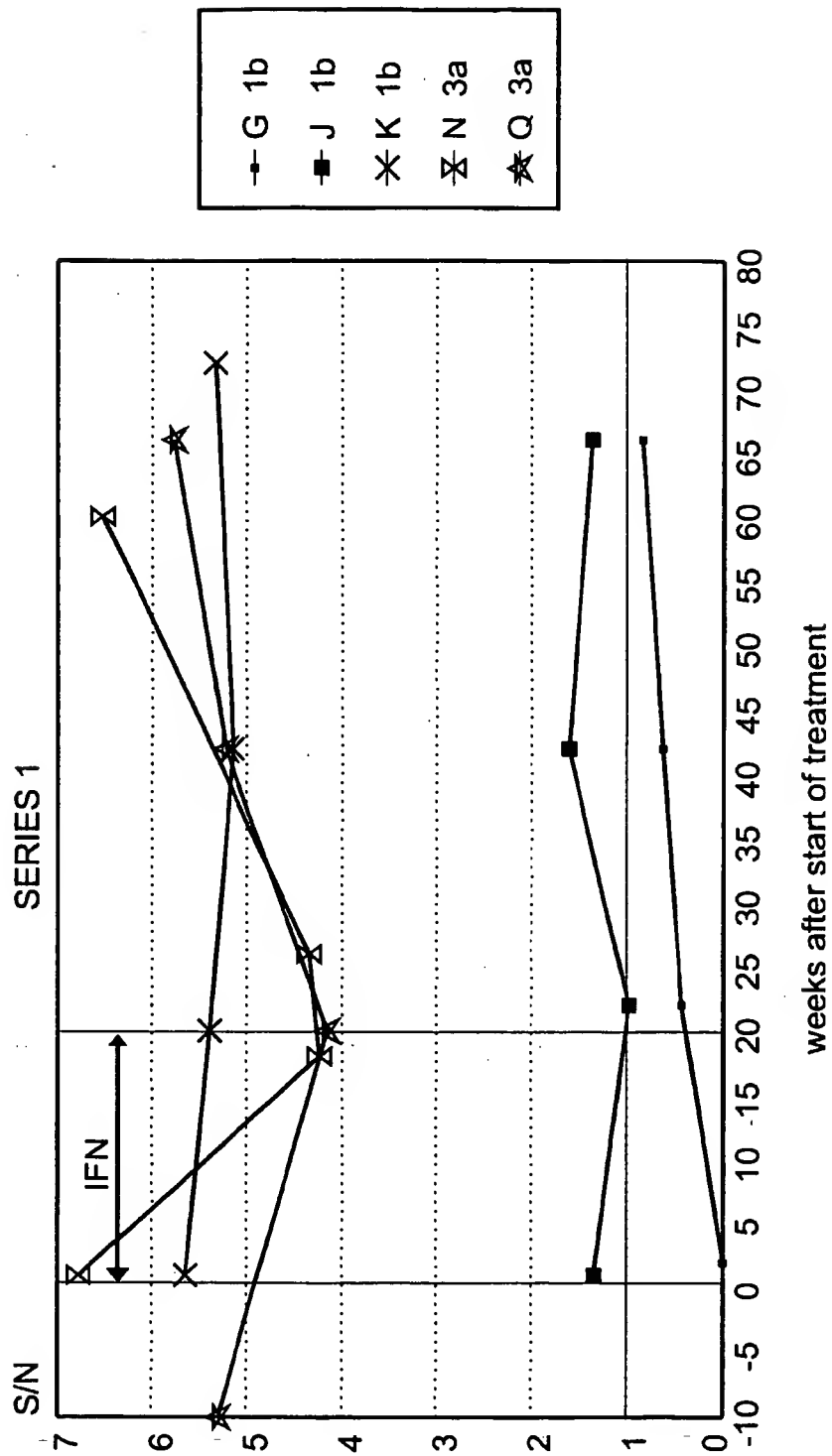


Figure 15

16 / 69

# Anti-E1 (epitope 1) levels in RESPONDERS to IFN treatment

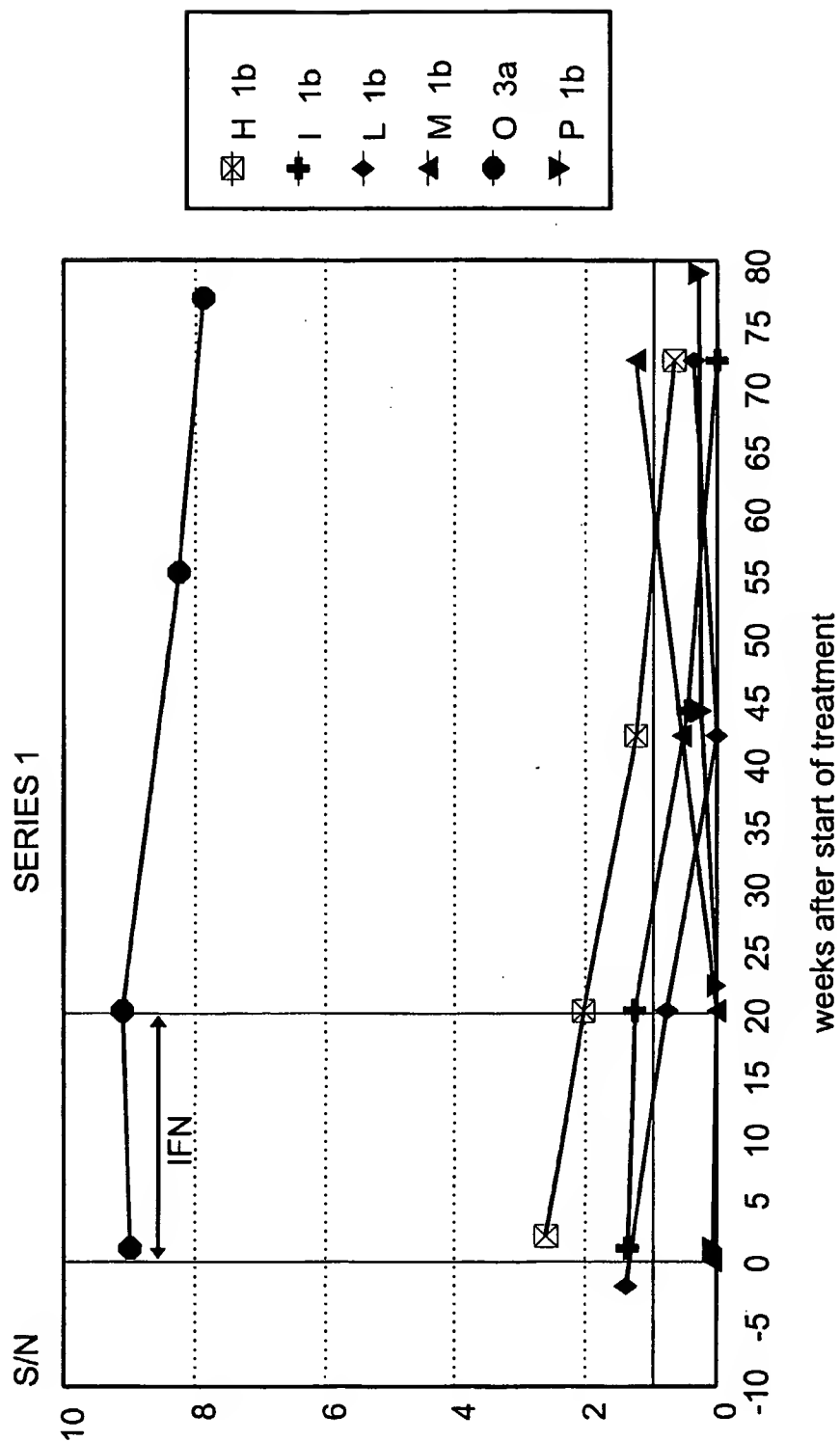


Figure 16



17 / 69

# Anti-E1 (epitope 2) levels in NON-RESPONDERS to IFN treatment

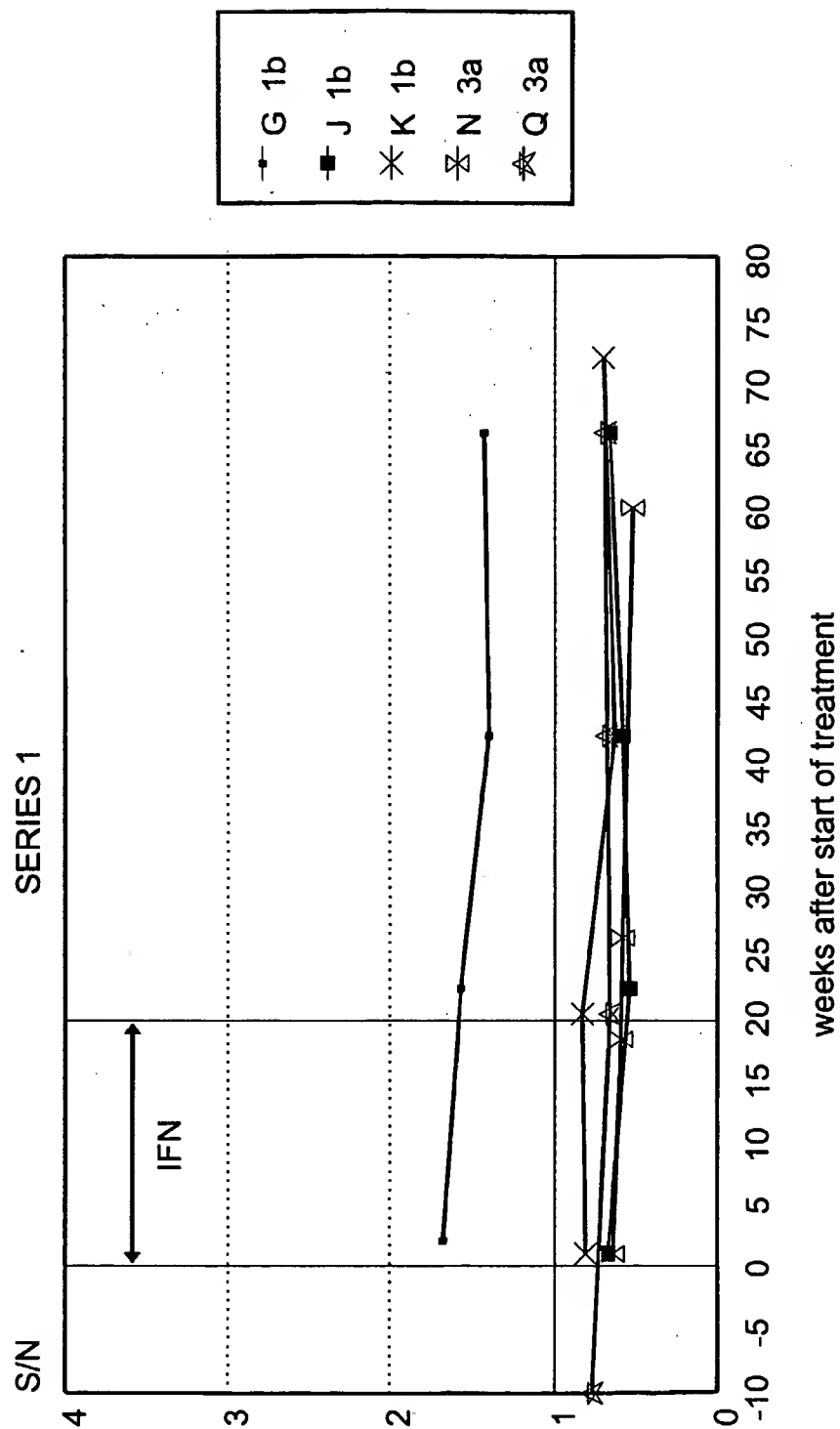


Figure 17

18 / 69

# Anti-E1 (epitope 2) levels in RESPONDERS to IFN treatment

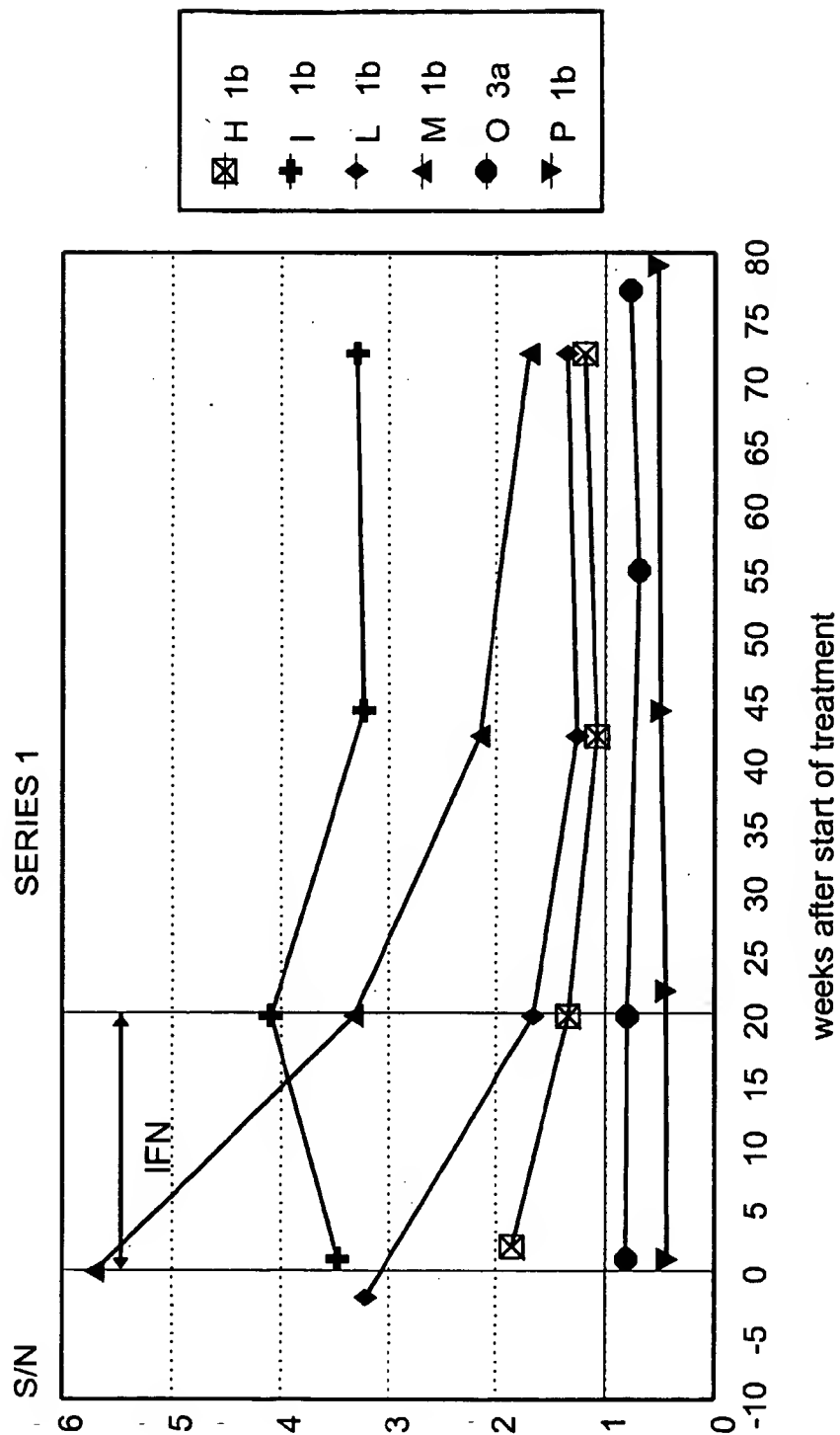


Figure 18

19 / 69

# Competition of reactivity of anti-E2 Mabs with peptides

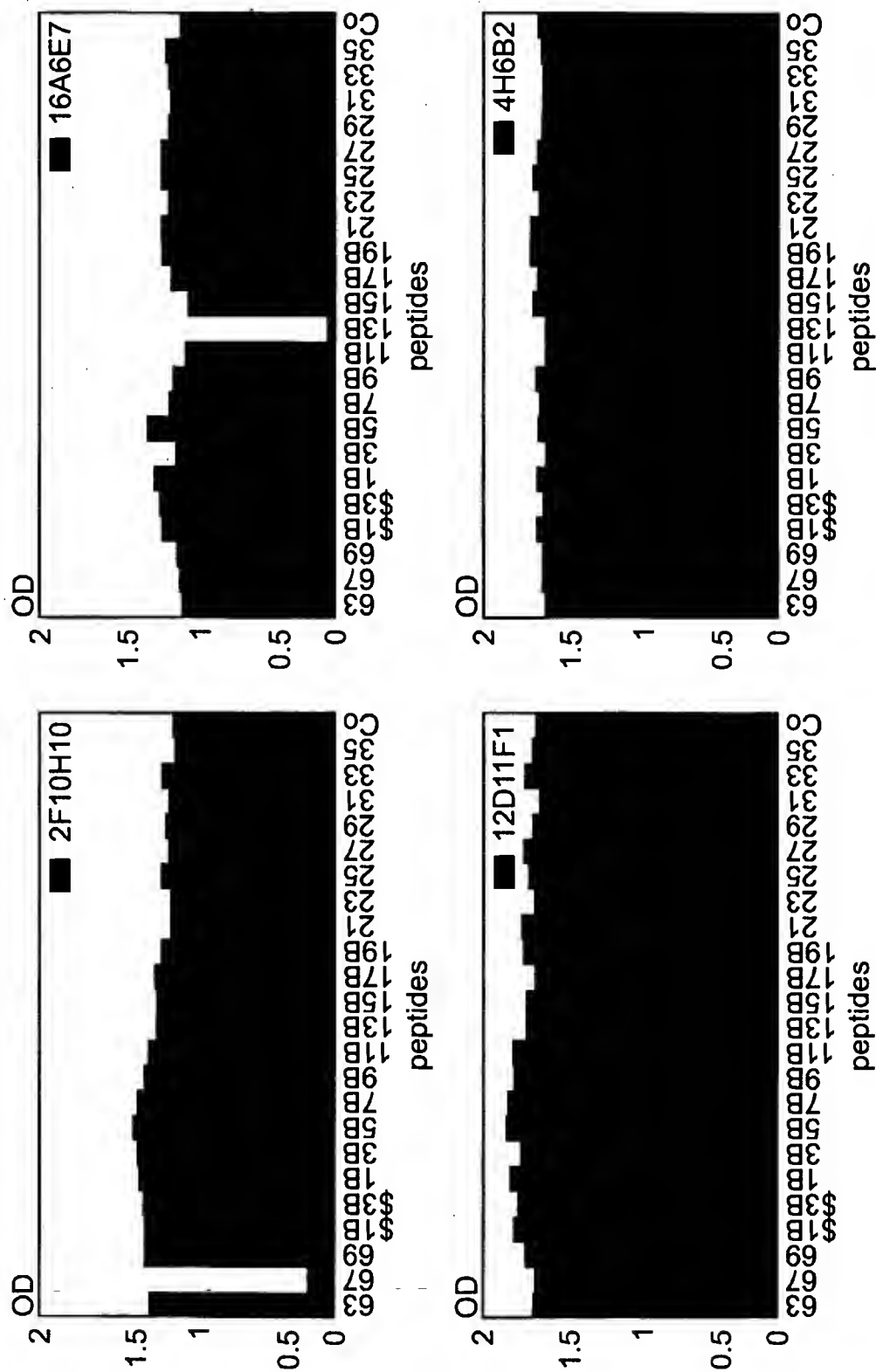


Figure 19

20 / 69

# Human anti-E2 reactivity competed with peptides

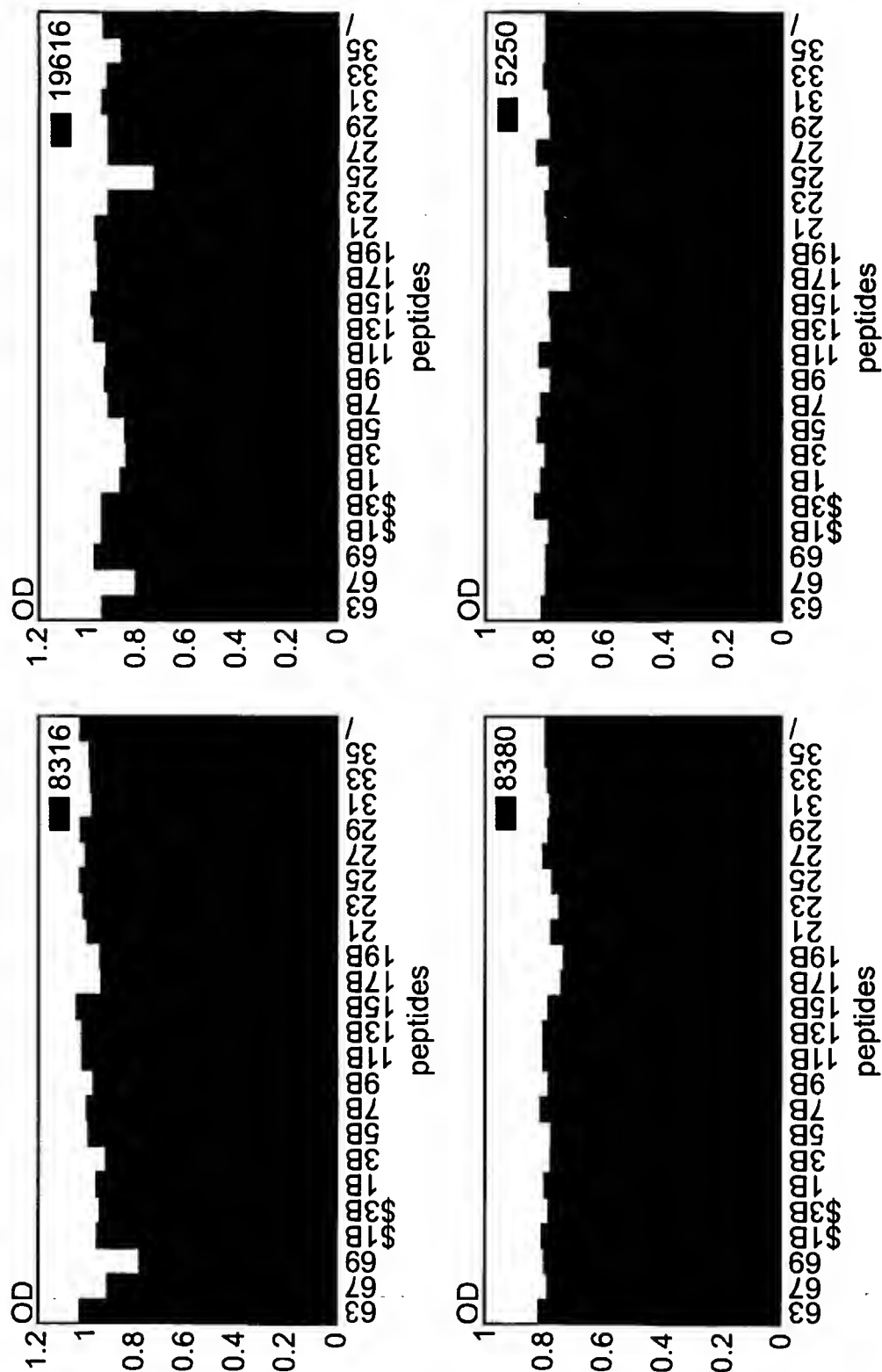


Figure 20

21 / 69

5' GGCATGCAAGCTTAATTAATT 3' (SEQ ID NO 1)  
3' ACGTCCGTACGTTCTGAATTAATTAATCGA 5' (SEQ ID NO 94)

5' CCGGGGAGGCCTGCACGTGATCGAGGGCAGACACCATCACCACCATCACTAATAGTTA  
ATTAAGTGCA 3' (SEQ ID NO 2)  
3' CCTCCGGACGTGCACTAGCTCCCGTCTGTGGTAGTGGTGGTAGTGATTATCAATTAAT  
TG 5' (SEQ ID NO 95)

SEQ ID NO 3 (HCC19A)  
ATGCCCGGTTGCTCTTTCTCTATCTTCCTCTTGGCTTTACTGTCCTGTCTGACCATTCCA  
GCTTCCGCTTATGAGGTGCGCAACGTGTCCGGGATGTACCATGTCACGAACGACTGCTCC  
AACTCAAGCATTGTGTATGAGGCAGCGGACATGATCATGCACACCCCCGGGTGCGTGCCC  
TGCGTTCCGGGAGAACAACTCTTCCCGCTGCTGGGTAGCGCTCACCCCCACGCTCGCAGCT  
AGGAACGCCAGCGTCCCCACCACGACAATACGACGCCACGTGCGATTTGCTCGTTGGGGCG  
GCTGCTCTCTGTTCCGCTATGTACGTGGGGGATCTCTGCGGATCTGTCTTCCTCGTCTCC  
CAGCTGTTCAACATCTCGCCTCGCCGGCATGAGACGGTGCAGGACTGCAATTGCTCAATC  
TATCCCGGCCACATAACAGGTCACCGTATGGCTTGGGATATGATGATGAACTGGTCGCCT  
ACAACGGCCCTGGTGGTATCGCAGCTGCTCCGGATCCCACAAGCTGTCGTGGACATGGTG  
GCGGGGGCCATTGGGGAGTCCTGGCGGGCCTCGCCTACTATTCCATGGTGGGGAACCTGG  
GCTAAGGTTTTGATTGTGATGCTACTCTTTGCTCTCTAATAG

SEQ ID NO 5 (HCC110A)  
ATGTTGGGTAAGGTCATCGATACCCTTACATGCGGCTTCGCCGACCTCGTGGGGTACATT  
CCGCTCGTCGGCGCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTT  
CTGGAGGACGGCGTGAACATATGCAACAGGGAATTTGCCCGGTTGCTCTTTCTCTATCTTC  
CTCTTGGCTTTGCTGTCTGTCTGACCGTTCCAGCTTCCGCTTATGAAGTGCGCAACGTG  
TCCGGGATGTACCATGTCACGAACGACTGCTCCAACCTCAAGCATTGTGTATGAGGCAGCG  
GACATGATCATGCACACCCCCGGGTGCGTGCCCTGCGTTCCGGGAGAACAACTCTTCCCGC  
TGCTGGGTAGCGCTCACCCCCACGCTCGCAGCTAGGAACGCCAGCGTCCCCACCACGACA  
ATACGACGCCACGTGCGATTTGCTCGTTGGGGCGGCTGCTTTCTGTTCCGCTATGTACGTG  
GGGGACCTCTGCGGATCTGTCTTCCTCGTCTCCAGCTGTTCAACATCTCGCCTCGCCGG  
CATGAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTACCGT  
ATGGCTTGGGATATGATGATGAACTGGTCGCCTACAACGGCCCTGGTGGTATCGCAGCTG  
CTCCGGATCCCACAAGCTGTCGTGGACATGGTGGCGGGGGCCATTGGGGAGTCCTGGCG  
GGTCTCGCCTACTATTCCATGGTGGGGAACCTGGGCTAAGGTTTTGATTGTGATGCTACTC  
TTTGCTCCCTAATAG

SEQ ID NO 7 (HCC111A)

Figure 21A

22 / 69

ATGTTGGGTAAGGTCATCGATACCCTTACGTGCGGCTTCGCCGACCTCATGGGGTACATT  
CCGCTCGTCGGCGCCCCCTAGGGGGTGCTGCCAGAGCCCTGGCGCATGGCGTCCGGGT  
CTGGAAGACGGCGTGAACATATGCAACAGGGAATTTGCCTGGTTGCTCTTTCTCTATCTTC  
CTCTTGGCTTTACTGTCTGTCTGACCATTCCAGCTTCCGCTTATGAGGTGCGCAACGTG  
TCCGGGATGTACCATGTACGAACGACTGCTCCAACCTCAAGCATTGTGTATGAGGCAGCG  
GACATGATCATGCACACCCCCGGGTGCGTGCCCTGCGTTCCGGAGAACAACCTCTTCCCGC  
TGCTGGGTAGCGCTCACCCCCACGCTCGCAGCTAGGAACGCCAGCGTCCCCACTACGACA  
ATACGACGCCACGTGATTTGCTCGTTGGGGCGGCTGCTTTCTGTTCCGCTATGTACGTG  
GGGGATCTCTGCGGATCTGTCTTCCCTCGTCTCCAGCTGTTCCACCATCTCGCCTCGCCGG  
CATGAGACGGTGCAAGACTGCAATTGCTCAATCTATCCCGGCCACATAACAGGTCACCGT  
ATGGCTTGGGATATGATGATGAACTGGTAATAG

SEQ ID NO 9 (HCC112A)

ATGCCCGGTTGCTCTTTCTCTATCTTCCCTCTTGGCCCTGCTGTCCTGTCTGACCATACCA  
GCTTCCGCTTATGAAGTGCGCAACGTGTCCGGGGTGTACCATGTACGAACGACTGCTCC  
AACTCAAGCATAGTGTATGAGGCAGCGGACATGATCATGCACACCCCCGGGTGCGTGCCC  
TGCGTTCCGGGAGGGCAACTCCTCCCGTTGCTGGGTGGCGCTCACTCCACGCTCGCGGCC  
AGGAACGCCAGCGTCCCCACAACGACAATACGACGCCACGTGATTTGCTCGTTGGGGCT  
GCTGCTTTCTGTTCCGCTATGTACGTGGGGGATCTCTGCGGATCTGTTTTCTTTGTTTCC  
CAGCTGTTACCTTCTCACCTCGCCGGCATCAAACAGTACAGGACTGCAACTGCTCAATC  
TATCCCGGCCATGTATCAGGTCACCGCATGGCTTGGGATATGATGATGAACTGGTCCTAA  
TAG

SEQ ID NO 11 (HCC113A)

ATGTCCGGTTGCTCTTTCTCTATCTTCCCTCTTGGCCCTGCTGTCCTGTCTGACCATACCA  
GCTTCCGCTTATGAAGTGCGCAACGTGTCCGGGGTGTACCATGTACGAACGACTGCTCC  
AACTCAAGCATAGTGTATGAGGCAGCGGACATGATCATGCACACCCCCGGGTGCGTGCCC  
TGCGTTCCGGGAGGGCAACTCCTCCCGTTGCTGGGTGGCGCTCACTCCACGCTCGCGGCC  
AGGAACGCCAGCGTCCCCACAACGACAATACGACGCCACGTGATTTGCTCGTTGGGGCT  
GCTGCTTTCTGTTCCGCTATGTACGTGGGGGATCTCTGCGGATCTGTTTTCTTTGTTTCC  
CAGCTGTTACCTTCTCACCTCGCCGGCATCAAACAGTACAGGACTGCAACTGCTCAATC  
TATCCCGGCCATGTATCAGGTCACCGCATGGCTTGGGATATGATGATGAACTGGTAATAG

SEQ ID NO 13 (HCC117A)

ATGCTGGGTAAGGCCATCGATACCCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACATT  
CCGCTCGTCGGCGCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGT  
CTGGAAGACGGCGTGAACATATGCAACAGGGAATTTGCCTGGTTGCTCTTTCTCTATCTTC  
CTCTTGGCTTTACTGTCTGTCTAACCATTCCAGCTTCCGCTTACGAGGTGCGCAACGTG  
TCCGGGATGTACCATGTACGAACGACTGCTCCAACCTCAAGCATTGTGTATGAGGCAGCG  
GACATGATCATGCACACCCCCGGGTGCGTGCCCTGCGTTCCGGAGAACAACCTCTTCCCGC  
TGCTGGGTAGCGCTCACCCCCACGCTCGCGGCTAGGAACGCCAGCATCCCCACTACAACA

Figure 21B

23 / 69

ATACGACGCCACGTCGATTTGCTCGTTGGGGCGGCTGCTTTCTGTTCCGCTATGTACGTG  
GGGGATCTCTGCGGATCTGTCTTCCTCGTCTCCAGCTGTTCAACATCTCGCCTCGCCGG  
CATGAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTCACCGT  
ATGGCTTGGGATATGATGATGAAGTGGTACTAATAG

SEQ ID NO 15 (HCP<sub>r</sub>51)

ATGCCCCGGTTGCTCTTTCTCTATCTT

SEQ ID NO 16 (HCP<sub>r</sub>52)

ATGTTGGGTAAGGTCATCGATACCCCT

SEQ ID NO 17 (HCP<sub>r</sub>53)

CTATTAGGACCAGTTCATCATCATATCCCA

SEQ ID NO 18 (HCP<sub>r</sub>54)

CTATTACCAGTTCATCATCATATCCCA

SEQ ID NO 19 (HCP<sub>r</sub>107)

ATACGACGCCACGTCGATTCCCAGCTGTTCAACATC

SEQ ID NO 20 (HCP<sub>r</sub>108)

GATGGTGAACAGCTGGGAATCGACGTGGCGTCGTAT

SEQ ID NO 21 (HCC137)

ATGTTGGGTAAGGTCATCGATACCCTTACATGCGGCTTCGCCGACCTCGTGGGGTACATT  
CCGCTCGTCGGCGCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTT  
CTGGAGGACGGCGTGAAGTATGCAACAGGGAATTTGCCCGGTTGCTCTTTCTCTATCTTC  
CTCTTGGCTTTGCTGTCTGTCTGACCGTTCCAGCTTCCGCTTATGAAGTGCACAACGTG  
TCCGGGATGTACCATGTCACGAACGACTGCTCCAAGTCAAGCATTGTGTATGAGGCAGCG  
GACATGATCATGCACACCCCCGGGTGCGTGCCCTGCGTTCCGGGAGAACAAGTCTTCCCGC  
TGCTGGGTAGCGCTCACCCCCACGCTCGCAGCTAGGAACGCCAGCGTCCCCACCACGACA  
ATACGACGCCACGTCGATTCCCAGCTGTTCAACATCTCGCCTCGCCGGCATGAGACGGTG  
CAGGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTCACCGTATGGCTTGGGAT  
ATGATGATGAAGTGGTGCCTACAACGGCCCTGGTGGTATCGCAGCTGCTCCGGATCCCA  
CAAGCTGTCTGGACATGGTGGCGGGGGCCATTGGGGAGTCTGGCGGGTCTCGCTAC  
TATTCCATGGTGGGGAAGTGGGCTAAGGTTTTGATTGTGATGCTACTCTTGCTCCCTAA  
TAG

SEQ ID NO 23 (HCC138)

ATGTTGGGTAAGGTCATCGATACCCTTACATGCGGCTTCGCCGACCTCGTGGGGTACATT  
CCGCTCGTCGGCGCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTT  
CTGGAGGACGGCGTGAAGTATGCAACAGGGAATTTGCCCGGTTGCTCTTTCTCTATCTTC  
CTCTTGGCTTTGCTGTCTGTCTGACCGTTCCAGCTTCCGCTTATGAAGTGCACAACGTG  
TCCGGGATGTACCATGTCACGAACGACTGCTCCAAGTCAAGCATTGTGTATGAGGCAGCG

Figure 21C

24 / 69

GACATGATCATGCACACCCCCGGGTGCGTGCCCTGCGTTTCGGGAGAACAACTCTTCCCGC  
TGCTGGGTAGCGCTCACCCCCACGCTCGCAGCTAGGAACGCCAGCGTCCCCACCACGACA  
ATACGACGCCACGTGATTCCCAGCTGTTCAACATCTCGCCTCGCCGGCATGAGACGGTG  
CAGGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTCACCGTATGGCTTGGGAT  
ATGATGATGAACTGGTAATAG

SEQ ID NO 25 (HCC139)

ATGTTGGGTAAGGTCATCGATACCCTTACATGCGGCTTCGCCGACCTCGTGGGGTACATT  
CCGCTCGTCGGCGCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTT  
CTGGAGGACGGCGTGAACATATGCAACAGGGAATTTGCCCGGTTGCTCTTTCTCTATCTTC  
CTCTTGGCTTTGCTGTCTGTCTGACCGTTCCAGCTTCCGCTTATGAAGTGCGCAACGTG  
TCCGGGATGTACCATGTACGAACGACTGCTCCAACCTCAAGCATTGTGTATGAGGCAGCG  
GACATGATCATGCACACCCCCGGGTGCGTGCCCTGCGTTTCGGGAGAACAACTCTTCCCGC  
TGCTGGGTAGCGCTCACCCCCACGCTCGCAGCTAGGAACGCCAGCGTCCCCACCACGACA  
ATACGACGCCACGTGATTCCCAGCTGTTCAACATCTCGCCTCGCCGGCATGAGACGGTG  
CAGGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTCACCGTATGGCTTGGGAT  
ATGATGATGAACTGGTCGCCTACAACGGCCCTGGTGGTATCGCAGCTGCTCCGGATCCTC  
TAATAG

SEQ ID NO 27 (HCC140)

ATGTTGGGTAAGGTCATCGATACCCTTACATGCGGCTTCGCCGACCTCGTGGGGTACATT  
CCGCTCGTCGGCGCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTT  
CTGGAGGACGGCGTGAACATATGCAACAGGGAATTTGCCCGGTTGCTCTTTCTCTATCTTC  
CTCTTGGCTTTGCTGTCTGTCTGACCGTTCCAGCTTCCGCTTATGAAGTGCGCAACGTG  
TCCGGGATGTACCATGTACGAACGACTGCTCCAACCTCAAGCATTGTGTATGAGGCAGCG  
GACATGATCATGCACACCCCCGGGTGCGTGCCCTGCGTTTCGGGAGAACAACTCTTCCCGC  
TGCTGGGTAGCGCTCACCCCCACGCTCGCAGCTAGGAACGCCAGCGTCCCCACCACGACA  
ATACGACGCCACGTGATTCCCAGCTGTTCAACATCTCGCCTCGCCGGCATGAGACGGTG  
CAGGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTCACCGTATGGCTTGGGAT  
ATGATGATGAACTGGTCGCCTACAACGGCCCTGGTGGTATCGCAGCTGCTCCGGATCGTG  
ATCGAGGGCAGACACCATCACCACCATCACTAATAG

SEQ ID NO 29 (HCC162)

ATGGGTAAGGTCATCGATACCCTTACGTGCGGATTGCCGATCTCATGGGGTACATCCCG  
CTCGTCGGCGCTCCCGTAGGAGGCGTCGCAAGAGCCCTTGCGCATGGCGTGAGGGCCCTT  
GAAGACGGGATAAATTTGCAACAGGGAATTTGCCCGGTTGCTCCTTTTCTATTTTCCTT  
CTCGCTCTGTTCTTTGCTTAATTCATCCAGCAGCTAGTCTAGAGTGCGGAATACGTCT  
GGCCTCTATGTCCTTACCAACGACTGTTCCAATAGCAGTATTGTGTACGAGGCCGATGAC  
GTTATTCTGCACACACCCGGCTGCATACCTTGTGTCCAGGACGGCAATACATCCACGTGC  
TGGACCCCACTGACACCTACAGTGGCAGTCAAGTACGTGGAGCAACCACCGCTTCGATA  
CGCAGTCATGTGGACCTATTAGTGGGCGCGGCCACGATGTGCTCTGCGCTCTACGTGGGT  
GACATGTGTGGGGCTGTCTTCCTCGTGGGACAAGCCTTCACGTTACAGACCTCGTCGCCAT

Figure 21D



25 / 69

CAAACGGTCCAGACCTGTAAGTCTGCTCGCTGTACCCAGGCCATCTTTCAGGACATCGAATG  
GCTTGGGATATGATGATGAACTGGTAATAG

SEQ ID NO 31 (HCC163)

ATGGGTAAGGTCATCGATACCCTAACGTGCGGATTGCGCGATCTCATGGGGTATATCCCG  
CTCGTAGGCGGGCCCCATTGGGGGCGTCGCAAGGGCTCTCGCACACGGTGTGAGGGTCCTT  
GAGGACGGGGTAACTATGCAACAGGGAATTTACCCGGTTGCTCTTCTCTATCTTTATT  
CTTGCTCTTCTCTCGTGTCTGACCGTTCCGGCCTCTGCAGTTCCTACCGAAATGCCTCT  
GGGATTTATCATGTTACCAATGATTGCCCAAACCTCTTCATAGTCTATGAGGCAGATAAC  
CTGATCCTACACGCACCTGGTTGCGTGCCTTGTGTCTATGACAGGTAATGTGAGTAGATGC  
TGGGTCCAAATTACCCCTACACTGTCAGCCCCGAGCCTCGGAGCAGTCACGGCTCCTCTT  
CGGAGAGCCGTTGACTACCTAGCGGGAGGGGCTGCCCTCTGCTCCGCGTTATACGTAGGA  
GACGCGTGTGGGGCACTATTCTTGGTAGGCCAAATGTTACCTATAGGCCTCGCCAGCAC  
GCTACGGTGCAGAACTGCAACTGTTCCATTTACAGTGGCCATGTTACCGGCCACCGGATG  
GCATGGGATATGATGATGAACTGGTAATAG

SEQ ID NO 33 (HCP109)

TGGGATATGATGATGAACTGGTC

SEQ ID NO 34 (HCP72)

CTATTATGGTGGTAAKGCCARCARGAGCAGGAG

SEQ ID NO 35 (HCCL22A)

TGGGATATGATGATGAACTGGTCGCCTACAACGGCCCTGGTGGTATCGCAGCTGCTCCGG  
ATCCCACAAGCTGTCTGGACATGGTGGCGGGGGCCCATTTGGGGAGTCCTGGCGGGCCTC  
GCCTACTATTCCATGGTGGGGAACCTGGGCTAAGGTTTTGGTTGTGATGCTACTCTTTGCC  
GGCGTCGACGGGCATACCCGCGTGTCTAGGAGGGGCGAGCAGCCTCCGATACCAGGGGCCTT  
GTGTCCCTCTTTAGCCCCGGGTGCGGCTCAGAAAATCCAGCTCGTAAACACCAACGGCAGT  
TGGCACATCAACAGGACTGCCCTGAACTGCAACGACTCCCTCCAAACAGGGTTCTTTGCC  
GCACTATTCTACAAACACAAATTCAACTCGTCTGGATGCCAGAGCGCTTGGCCAGCTGT  
CGCTCCATCGACAAGTTCTGCTCAGGGGTGGGGTCCCCTCACTTAACTGAGCCTAACAGC  
TCGGACCAGAGGGCCCTACTGCTGGCACTACGCGCCTCGACCGTGTGGTATTGTACCCGCG  
TCTCAGGTGTGCGGTCCAGTGTATTGCTTCACCCCGAGCCCTGTTGTGGTGGGGACGACC  
GATCGGTTTGGTGTCCCCACGTATAACTGGGGGGCGAACGACTCGGATGTGCTGATTCTC  
AACAACACGCGGCCGCGCGAGGCAACTGGTTGCGGCTGTACATGGATGAATGGCACTGGG  
TTCACCAAGACGTGTGGGGGCCCCCGTGCAACATCGGGGGGGCCGGCAACAACACCTTG  
ACCTGCCCCACTGACTGTTTTCGGAAGCACCCCGAGGCCACCTACGCCAGATGCGGTTCT  
GGGCCCTGGCTGACACCTAGGTGTATGGTTCACTACCCATATAGGCTCTGGCACTACCCC  
TGCACTGTCAACTTCACCATCTTCAAGGTTAGGATGTACGTGGGGGGCGTGGAGCACAGG  
TTCGAAGCCGCATGCAATTGGAATCGAGGAGAGCGTTGTGACTTGGAGGACAGGGATAGA  
TCAGAGCTTAGCCCGCTGCTGCTGTCTACAACAGAGTGGCAGATACTGCCCTGTTCTTC

Figure 21E

26 / 69

ACCACCCTGCCGGCCCTATCCACCGGCCTGATCCACCTCCATCAGAACATCGTGGACGTG  
CAATACCTGTACGGTGTAGGGTCGGCGGTTGTCTCCCTTGTCATCAAATGGGAGTATGTC  
CTGTTGCTCTTCCTTCTCCTGGCAGACGCGCGCATCTGCGCCTGCTTATGGATGATGCTG  
CTGATAGCTCAAGCTGAGGCCGCCTTAGAGAACCTGGTGGTCCTCAATGCGGCGGCCGTG  
GCCGGGGCGCATGGCACTCTTTCCTTCTTGTGTTCTTCTGTGCTGCCTGGTACATCAAG  
GGCAGGCTGGTCCCTGGTGGGCATACGCCTTCTATGGCGTGTGGCCGCTGCTCCTGCTT  
CTGCTGGCCTTACCACCACGAGCTTATGCCTAGTAA

SEQ ID NO 37 (HCC141)

GATCCACACAAGCTGTCGTGGACATGGTGGCGGGGGCCATTGGGGAGTCCTGGCGGGCCT  
CGCCTACTATTCCATGGTGGGGAACCTGGGCTAAGGTTTTGGTTGTGATGCTACTCTTTGC  
CGGCGTCGACGGGCATACCCGCGTGTGAGGAGGGGCAGCAGCCTCCGATACCAGGGGCCT  
TGTGTCCCTCTTTAGCCCCGGGTGCGCTCAGAAAATCCAGCTCGTAAACACCAACGGCAG  
TTGGCACATCAACAGGACTGCCCTGAACTGCAACGACTCCCTCCAAACAGGGTTCTTTGC  
CGCACTATTCTACAAACACAAATTCAACTCGTCTGGATGCCAGAGCGCTTGGCCAGCTG  
TCGCTCCATCGACAAGTTCGCTCAGGGGTGGGGTCCCCTCACTTACACTGAGCCTAACAG  
CTCGGACCAGAGGCCCTACTGCTGGCACTACGCGCCTCGACCGTGTGGTATTGTACCCGC  
GTCTCAGGTGTGCGGTCCAGTGTATTGCTTACCCCGAGCCCTGTTGTGGTGGGGACGAC  
CGATCGGTTTTGGTGTCCCCACGTATAACTGGGGGGCGAACGACTCGGATGTGCTGATTCT  
CAACAACACGCGGCCGCGCGAGGCAACTGGTTCGGCTGTACATGGATGAATGGCACTGG  
GTTACCAAGACGTGTGGGGGCCCCCGTGCAACATCGGGGGGGCCGGCAACAACACCTT  
GACCTGCCCCACTGACTGTTTTCGGAAGCACCCCGAGGCCACCTACGCCAGATGCGGTTT  
TGGGCCCTGGCTGACACCTAGGTGTATGGTTCATTACCCATATAGGCTCTGGCACTACCC  
CTGCACTGTCAACTTCACCATCTTCAAGGTTAGGATGTACGTGGGGGGCGTGGAGCACAG  
GTTTGAAGCCGCATGCAATTGGAAGTGGAGAGAGCGTTGTGACTTGGAGGACAGGGATAG  
ATCAGAGCTTAGCCCGCTGCTGCTGTCTACAACAGAGTGGCAGAGTGGCAGAGCTTAATT  
AATTAG

SEQ ID NO 39 (HCC142)

GATCCACACAAGCTGTCGTGGACATGGTGGCGGGGGCCATTGGGGAGTCCTGGCGGGCCT  
CGCCTACTATTCCATGGTGGGGAACCTGGGCTAAGGTTTTGGTTGTGATGCTACTCTTTGC  
CGGCGTCGACGGGCATACCCGCGTGTGAGGAGGGGCAGCAGCCTCCGATACCAGGGGCCT  
TGTGTCCCTCTTTAGCCCCGGGTGCGCTCAGAAAATCCAGCTCGTAAACACCAACGGCAG  
TTGGCACATCAACAGGACTGCCCTGAACTGCAACGACTCCCTCCAAACAGGGTTCTTTGC  
CGCACTATTCTACAAACACAAATTCAACTCGTCTGGATGCCAGAGCGCTTGGCCAGCTG  
TCGCTCCATCGACAAGTTCGCTCAGGGGTGGGGTCCCCTCACTTACACTGAGCCTAACAG  
CTCGGACCAGAGGCCCTACTGCTGGCACTACGCGCCTCGACCGTGTGGTATTGTACCCGC  
GTCTCAGGTGTGCGGTCCAGTGTATTGCTTACCCCGAGCCCTGTTGTGGTGGGGACGAC  
CGATCGGTTTTGGTGTCCCCACGTATAACTGGGGGGCGAACGACTCGGATGTGCTGATTCT  
CAACAACACGCGGCCGCGCGAGGCAACTGGTTCGGCTGTACATGGATGAATGGCACTGG  
GTTACCAAGACGTGTGGGGGCCCCCGTGCAACATCGGGGGGGCCGGCAACAACACCTT

Figure 21F

27 / 69

GACCTGCCCCACTGACTGTTTTTCGGAAGCACCCCGAGGCCACCTACGCCAGATGCGGTTC  
TGGGCCCTGGCTGACACCTAGGTGTATGGTTCATTACCCATATAGGCTCTGGCACTACCC  
CTGCACTGTCAACTTCACCATCTTCAAGGTTAGGATGTACGTGGGGGGCGTGGAGCACAG  
GTTTGAAGCCGCATGCAATTGGACTCGAGGAGAGCGTTGTGACTTGGAGGACAGGGATAG  
ATCAGAGCTTAGCCCGCTGCTGCTGTCTACAACAGGTGATCGAGGGCAGACACCATCACC  
ACCATCACTAATAG

SEQ ID NO 41 (HCC143)

ATGGTGGGGAACTGGGCTAAGGTTTTGGTTGTGATGCTACTCTTTGCCGGCGTCGACGGG  
CATACCCGCGTGTGAGGAGGGGCAGCAGCCTCCGATACCAGGGGCCTTGTGTCCCTCTTT  
AGCCCCGGGTGGCTCAGAAAATCCAGCTCGTAAACACCAACGGCAGTTGGCACATCAAC  
AGGACTGCCCTGAACTGCAACGACTCCCTCCAAACAGGGTTCTTTGCCGCACTATTCTAC  
AAACACAAATTCAACTCGTCTGGATGCCAGAGCGCTTGGCCAGCTGTGCTCCATCGAC  
AAGTTCGCTCAGGGGTGGGGTCCCTCACTTACACTGAGCCTAACAGCTCGGACCAGAGG  
CCCTACTGCTGGCACTACGCGCCTCGACCGTGTGGTATTGTACCCGCGTCTCAGGTGTGC  
GGTCCAGTGTATTGCTTCACCCCGAGCCCTGTTGTGGTGGGGACGACCGATCGGTTTGGT  
GTCCCCACGTATAACTGGGGGGCGAACGACTCGGATGTGCTGATTCTCAACAACACGCGG  
CCGCCGCGAGGCAACTGGTTCGGCTGTACATGGATGAATGGCACTGGGTTACCAAGACG  
TGTGGGGGCCCCCGTGCAACATCGGGGGGGCGGCAACAACACCTTGACCTGCCCCACT  
GACTGTTTTCGGAAGCACCCCGAGGCCACCTACGCCAGATGCGGTTCTGGGCCCTGGCTG  
ACACCTAGGTGTATGGTTCATTACCCATATAGGCTCTGGCACTACCCCTGCACTGTCAAC  
TTCACCATCTTCAAGGTTAGGATGTACGTGGGGGGCGTGGAGCACAGGTTTGAAGCCGCA  
TGCAATTGGACTCGAGGAGAGCGTTGTGACTTGGAGGACAGGGATAGATCAGAGCTTAGC  
CCGCTGCTGCTGTCTACAACAGAGTGGCAGAGCTTAATTAATTAG

SEQ ID NO 43 (HCC144)

ATGGTGGGGAACTGGGCTAAGGTTTTGGTTGTGATGCTACTCTTTGCCGGCGTCGACGGG  
CATACCCGCGTGTGAGGAGGGGCAGCAGCCTCCGATACCAGGGGCCTTGTGTCCCTCTTT  
AGCCCCGGGTGGCTCAGAAAATCCAGCTCGTAAACACCAACGGCAGTTGGCACATCAAC  
AGGACTGCCCTGAACTGCAACGACTCCCTCCAAACAGGGTTCTTTGCCGCACTATTCTAC  
AAACACAAATTCAACTCGTCTGGATGCCAGAGCGCTTGGCCAGCTGTGCTCCATCGAC  
AAGTTCGCTCAGGGGTGGGGTCCCTCACTTACACTGAGCCTAACAGCTCGGACCAGAGG  
CCCTACTGCTGGCACTACGCGCCTCGACCGTGTGGTATTGTACCCGCGTCTCAGGTGTGC  
GGTCCAGTGTATTGCTTCACCCCGAGCCCTGTTGTGGTGGGGACGACCGATCGGTTTGGT  
GTCCCCACGTATAACTGGGGGGCGAACGACTCGGATGTGCTGATTCTCAACAACACGCGG  
CCGCCGCGAGGCAACTGGTTCGGCTGTACATGGATGAATGGCACTGGGTTACCAAGACG  
TGTGGGGGCCCCCGTGCAACATCGGGGGGGCGGCAACAACACCTTGACCTGCCCCACT  
GACTGTTTTCGGAAGCACCCCGAGGCCACCTACGCCAGATGCGGTTCTGGGCCCTGGCTG  
ACACCTAGGTGTATGGTTCATTACCCATATAGGCTCTGGCACTACCCCTGCACTGTCAAC  
TTCACCATCTTCAAGGTTAGGATGTACGTGGGGGGCGTGGAGCACAGGTTTGAAGCCGCA  
TGCAATTGGACTCGAGGAGAGCGTTGTGACTTGGAGGACAGGGATAGATCAGAGCTTAGC

Figure 21G

28 / 69

CCGCTGCTGCTGTCTACAACAGGTGATCGAGGGCAGACACCATCACCACCATCACTAATA  
G

SEQ ID NO 45 (HCCL64)

ATGGTGGCGGGGGCCATTGGGGAGTCCTGGCGGGCCTCGCCTACTATTCCATGGTGGGG  
AACTGGGCTAAGGTTTTGGTTGTGATGCTACTCTTTGCCGGCGTCGACGGGCATACCCGC  
GTGTCAGGAGGGGCAGCAGCCTCCGATACCAGGGGCCTTGTGTCCCTCTTTAGCCCCGGG  
TCGGCTCAGAAAATCCAGCTCGTAAACACCAACGGCAGTTGGCACATCAACAGGACTGCC  
CTGAACTGCAACGACTCCCTCCAAACAGGGTTCTTTGCCGCACTATTCTACAAACACAAA  
TTCAACTCGTCTGGATGCCCAGAGCGCTTGGCCAGCTGTGCTCCATCGACAAGTTGCT  
CAGGGGTGGGGTCCCCTCACTTACACTGAGCCTAACAGCTCGGACCAGAGGCCCTACTGC  
TGGCACTACGCGCCTCGACCGTGTGGTATTGTACCCGCGTCTCAGGTGTGCGGTCCAGTG  
TATTGCTTACCCCCGAGCCCTGTTGTGGTGGGGACGACCGATCGGTTTGGTGTCCCCACG  
TATAACTGGGGGGCGAACGACTCGGATGTGCTGATTCTCAACAACACGCGGCCGCCGCGA  
GGCAACTGGTTCCGGCTGTACATGGATGAATGGCACTGGGTTACCAAGACGTGTGGGGGC  
CCCCCGTGCAACATCGGGGGGGCCGGCAACAACACCTTGACCTGCCCCACTGACTGTTTT  
CGGAAGCACCCCCGAGGCCACCTACGCCAGATGCGGTTCTGGGCCCTGGCTGACACCTAGG  
TGTATGGTTCATTACCCATATAGGCTCTGGCACTACCCCTGCACTGTCAACTTCACCATC  
TTCAAGGTTAGGATGTACGTGGGGGGCGTGGAGCACAGGTTTGAAGCCGCATGCAATTGG  
ACTCGAGGAGAGCGTTGTGACTTGGAGGACAGGGATAGATCAGAGCTTAGCCCGCTGCTG  
CTGTCTACAACAGAGTGGCAGATACTGCCCTGTTCTTACCACCCTGCCGGCCCTATCC  
ACCGGCCTGATCCACCTCCATCAGAACATCGTGGACGTGCAATACCTGTACGGTGTAGGG  
TCGGCGGTTGTCTCCCTTGTCTCATCAATGGGAGTATGTCCTGTTGCTCTTCTCTCTG  
GCAGACGCGCGCATCTGCGCTGCTTATGGATGATGCTGCTGATAGCTCAAGCTGAGGCC  
GCCTTAGAGAACCTGGTGGTCCTCAATGCGGCGGCCGTGGCCGGGGCGCATGGCACTCTT  
TCCTTCTTGTGTTCTTCTGTGCTGCCTGGTACATCAAGGGCAGGCTGGTCCCTGGTGCG  
GCATACGCCTTCTATGGCGTGTGGCCGCTGCTCCTGCTTCTGCTGGCCTTACCACCACGA  
GCTTATGCCTAGTAA

SEQ ID NO 47 (HCC165)

AATTTGGGTAAGGTCATCGATACCCTTACATGCGGCTTCGCCGACCTCGTGGGGTACATT  
CCGCTCGTCGGCGCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTT  
CTGGAGGACGGCGTGAACATATGCAACAGGGAATTTGCCCGGTTGCTCTTTCTCTATCTTC  
CTCTTGGCTTTGCTGTCTGTCTGACCGTTCCAGCTTCCGCTTATGAAGTGCGCAACGTG  
TCCGGGATGTACCATGTACGAACGACTGCTCCAACCTCAAGCATTGTGTATGAGGCAGCG  
GACATGATCATGCACACCCCCGGGTGCGTGCCCTGCGTTCCGGGAGAACAACCTCTTCCCGC  
TGCTGGGTAGCGCTACCCCCACGCTCGCAGCTAGGAACGCCAGCGTCCCACCACGACA  
ATACGACGCCACGTGATTTGCTCGTTGGGGCGGCTGCTTTCTGTTCCGCTATGTACGTG  
GGGGACCTCTGCGGATCTGTCTTCTCGTCTCCAGCTGTTACCATCTCGCCTCGCCGG  
CATGAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTACCGT  
ATGGCTTGGGATATGATGATGAACTGGTCGCCTACAACGGCCCTGGTGGTATCGCAGCTG  
CTCCGGATCCCAAGCTGTCGTGGACATGGTGGCGGGGGCCATTGGGGAGTCCTGGCG

Figure 21H

29 / 69

GGCCTCGCCTACTATTCCATGGTGGGGAAGTGGGCTAAGGTTTTGGTTGTGATGCTACTC  
TTTGCCGGCGTTCGACGGGCATACCCGCGTGTGAGGAGGGGCAGC AGCCTCCGATACCAG  
GGGCCTTGTGTCCCTCTTTAGCCCCGGGTGCGCTCAGAAAATCCAGCTCGTAAACACCAA  
CGGCAGTTGGCACATCAACAGGACTGCCCTGAACTGCAACGACTCCCTCCAAACAGGGTT  
CTTTGCCGCACTATTCTACAAACACAAATTCAACTCGTCTGGATGCCAGAGCGCTTGGC  
CAGCTGTGCTCCATCGACAAGTTCGCTCAGGGGTGGGGTCCCCTCACTTACACTGAGCC  
TAACAGCTCGGACCAGAGGCCCTACTGCTGGCACTACGCGCTCGACCGTGTGGTATTGT  
ACCCGCGTCTCAGGTGTGCGGTCCAGTGTATTGCTTCACCCCGAGCCCTGTTGTGGTGGG  
GACGACCGATCGGTTTGGTGTCCCCACGTATAACTGGGGGGCGAACGACTCGGATGTGCT  
GATTCTCAACAACACGCGGCCCGCGAGGCAACTGGTTCGGCTGTACATGGATGAATGG  
CACTGGGTTACCAAGACGTGTGGGGGCCCCCGTGCAACATCGGGGGGGCGGCAACAA  
CACCTTGACCTGCCCCACTGACTGTTTTCGGAAGCACCCCGAGGCCACCTACGCCAGATG  
CGGTTCTGGGGCCCTGGCTGACACCTAGGTGTATGGTTCATTACCCATATAGGCTCTGGCA  
CTACCCCTGCACTGTCAACTTCACCATCTTCAAGGTTAGGATGTACGTGGGGGGCGTGGA  
GCACAGGTTCAAGCCGCATGCAATTGGAAGTTCGAGGAGAGCGTTGTGACTTGAGGACAG  
GGATAGATCAGAGCTTAGCCCGCTGCTGCTGTCTACAACAGAGTGGCAGATACTGCCCTG  
TTCCTTCACCACCCTGCCGGCCCTATCCACCGGCCTGATCCACCTCCATCAGAACATCGT  
GGACGTGCAATACCTGTACGGTGTAGGGTCGGCGGTTGTCTCCCTTGTGATCAAATGGGA  
GTATGTCCTGTTGCTCTTCTCTCTGCGCAGACGCGCGCATCTGCGCCTGCTTATGGAT  
GATGCTGCTGATAGCTCAAGCTGAGGCCGCTTAGAGAACCTGGTGGTCTCAATGCGGC  
GGCCGTGGCCGGGGCGCATGGCACTCTTTCCTTCTTGTGTTCTTCTGTGCTGCCTGGTA  
CATCAAGGGCAGGCTGGTCCCTGGTGGGCATACGCCCTTCTATGGCGTGTGGCCGCTGCT  
CCTGCTTCTGCTGGCCTTACCACCACGAGCTTATGCCTAGTAAGCTT

SEQ ID NO 49 (HCC166)

ATGAGCACGAATCCTAAACCTCAAAGAAAAACCAAACGTAACACCAACCGCCGCCACAG  
GACGTCAAGTTCCCGGGCGGTGGTCAGATCGTTGGTGGAGTTTACCTGTTGCCGCGCAGG  
GGCCCCAGGTTGGGTGTGCGCGCGACTAGGAAGACTTCGAGCGGTGCAACCTCGTGGG  
AGGCGACAACCTATCCCAAGGCTCGCCGACCCGAGGGTAGGGCCTGGGCTCAGCCCGGG  
TACCCTTGCCCCCTCTATGGCAATGAGGGCATGGGGTGGGCAGGATGGCTCCTGTACCC  
CGCGGCTCTCGGCCTAGTTGGGGCCCTACAGACCCCGGCGTAGGTGCGGTAATTTGGGT  
AAGGTCATCGATACCCTTACATGCGGCTTCGCCGACCTCGTGGGGTACATTCCGCTCGTC  
GGCGCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTTCTGGAGGAC  
GGCGTGAACATGCAACAGGGAATTTGCCCGGTTGCTCTTCTCTATCTTCTTGGCT  
TTGCTGTCTGTCTGACCGTTCCAGCTTCCGCTTATGAAGTGCAGCAACGTGTCCGGGATG  
TACCATGTACGAACGACTGCTCCAACCTCAAGCATTGTGTATGAGGCAGCGGACATGATC  
ATGCACACCCCGGGTGCGTGCCCTGCGTTCCGGGAGAACAACCTTCCCGCTGCTGGGTA  
GCGCTCACCCCCACGCTCGCAGCTAGGAACGCCAGCGTCCCCACCACGACAATACGACGC  
CACGTCGATTTGCTCGTTGGGGCGGCTGCTTCTGTTCCGCTATGTACGTGGGGGACCTC  
TGCGGATCTGTCTTCTCGTCTCCAGCTGTTACCATCTCGCCTCGCCGGCATGAGACG  
GTGCAGGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGTACCGTATGGCTTGG

Figure 21l

30 / 69

GATATGATGATGAACTGGTCGCCTACAACGGCCCTGGTGGTATCGCAGCTGCTCCGGATC  
CCACAAGCTGTCGTGGACATGGTGGCGGGGGCCATTGGGGAGTCCTGGCGGGCCTCGCC  
TACTATTCCATGGTGGGGAAC TGGGCTAAGGTTTTGGTTGTGATGCTACTCTTTGCCGGC  
GTCGACGGGCATACCCGCGTGT CAGGAGGGGCAGCAGCCTCCGATACCAGGGGCCTTGTC  
TCCCTCTTTAGCCCCGGGTGGGCTCAGAAAATCCAGCTCGTAAACACCAACGGCAGTTGG  
CACATCAACAGGACTGCCCTGAACTGCAACGACTCCCTCCAAACAGGGTTCTTTGCCGCA  
CTATTCTACAAACACAAATTCAACTCGTCTGGATGCCCAGAGCGCTTGGCCAGCTGTCGC  
TCCATCGACAAGTTCGCTCAGGGGTGGGGTCCCCTCACTTACACTGAGCCTAACAGCTCG  
GACCAGAGGGCCCTACTGCTGGCACTACGCGCCTCGACCGTGTGGTATTGTACCCGCGTCT  
CAGGTGTGCGGTCCAGTGTATTGCTTCACCCCGAGCCCTGTTGTGGTGGGGACGACCGAT  
CGGTTTGGTGTCCCCACGTATAACTGGGGGGCGAACGACTCGGATGTGCTGATTCTCAAC  
AACACGCGGGCCGCGCGAGGCAACTGGTTCGGCTGTACATGGATGAATGGCACTGGGTTC  
ACCAAGACGTGTGGGGGGCCCCCGTGCAACATCGGGGGGGCCGGCAACAACACCTTGACC  
TGCCCCACTGACTGTTTTCGGAAGCACCCCGAGGCCACCTACGCCAGATGCGGTTCTGGG  
CCCTGGCTGACACCTAGGTGTATGGTTCATTACCCATATAGGCTCTGGCACTACCCCTGC  
ACTGTCAACTTCACCATCTTCAAGTTAGGATGTACGTGGGGGGCGTGGAGCACAGGTTTC  
GAAGCCGCATGCAATTGGA CTGAGGAGAGCGTTGTGACTTGGAGGACAGGGATAGATCA  
GAGCTTAGCCCGCTGCTGCTGTCTACAACAGAGTGGCAGATACTGCCCTGTTCTTCACC  
ACCCTGCCGGCCCTATCCACCGGCCTGATCCACCTCCATCAGAACATCGTGGACGTGCAA  
TACCTGTACGGTGTAGGGTCGGCGGTTGTCTCCCTTGT CATCAAATGGGAGTATGTCCTG  
TTGCTCTTCCTTCTCCTGGCAGACGCGCGCATCTGCGCCTGCTTATGGATGATGCTGCTG  
ATAGCTCAAGCTGAGGCCGCCCTTAGAGAACCTGGTGGTCCTCAATGCGGCGGGCCGTGGCC  
GGGGCGCATGGCACTCTTTCTTCCTTGTGTTCTTCTGTGCTGCCTGGTACATCAAGGGC  
AGGCTGGTCCCTGGTGGGCATACGCCTTCTATGGCGTGTGGCCGCTGCTCCTGCTTCTG  
CTGGCCTTACCACCACGAGCTTATGCCTAGTAA

Figure 21J

31 / 69

OD measured at 450nm  
construct

Fraction	Volume	dilution	39 type 1b	40 type 1b	62 type 3a	63 type 5a
Start	23ml	1/20	2.517	1.954	1.426	1.142
Flow through	23ml	1/20	0.087	0.085	0.176	0.120
1	0.4ml	1/200	0.102	0.051	0.048	0.050
2			0.396	0.550	0.090	0.067
3			2.627	2.603	2.481	2.372
4			3	2.967	3	2.694
5			3	2.810	2.640	2.154
6			2.694	2.499	1.359	1.561
7			2.408	2.481	0.347	1.390
8			2.176	1.970	1.624	0.865
9			1.461	1.422	0.887	0.604
10			1.286	0.926	0.543	0.519
11			0.981	0.781	0.294	0.294
12			0.812	0.650	0.249	0.199
13			0.373	0.432	0.239	0.209
14			0.653	0.371	0.145	0.184
15			0.441	0.348	0.151	0.151
16			0.321	0.374	0.098	0.106
17			0.525	0.186	0.099	0.108
18			0.351	0.171	0.083	0.090
19			0.192	0.164	0.084	0.087

Figure 22

32 / 69

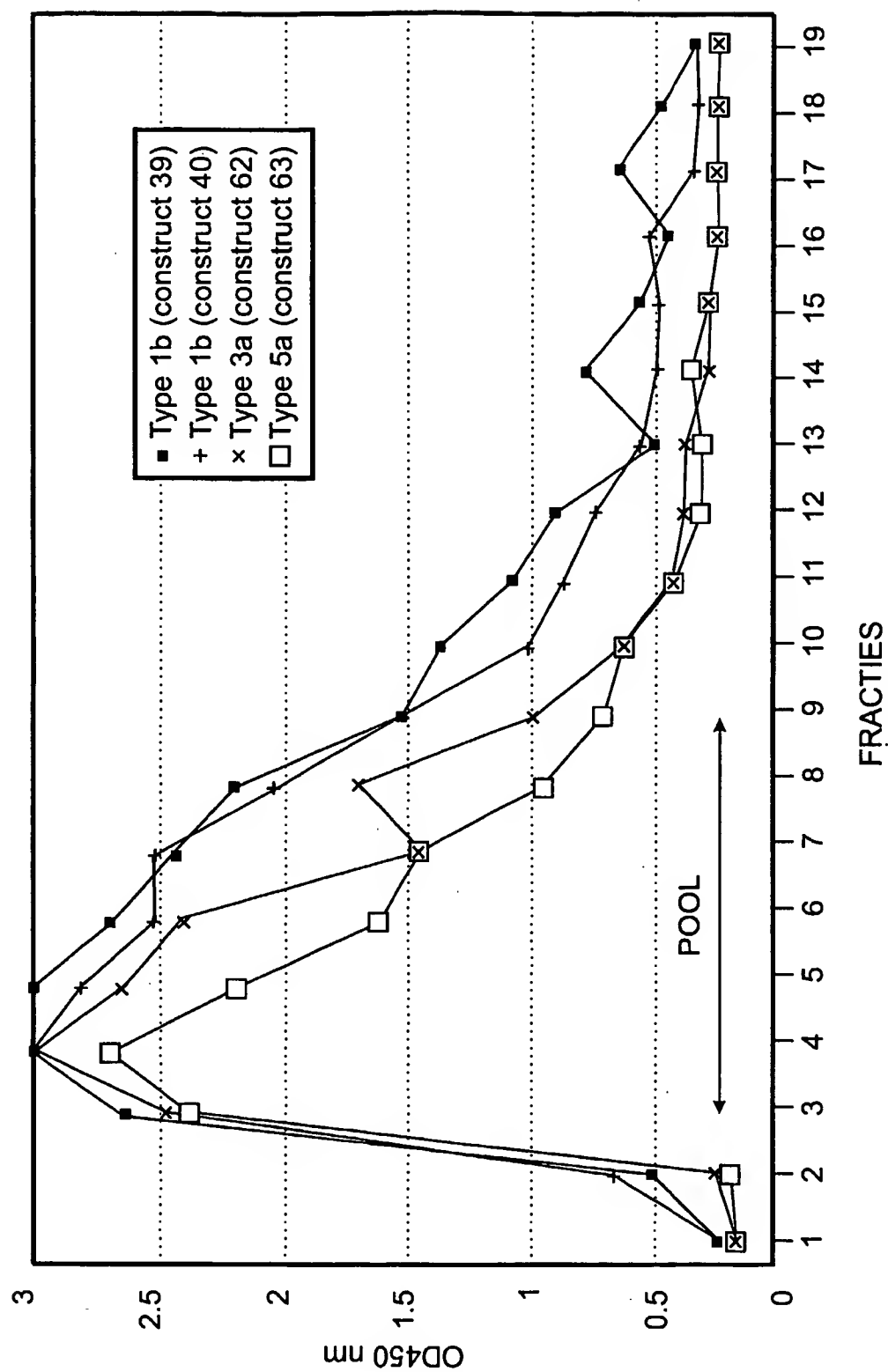


Figure 23



33 / 69

OD measured at 450nm  
construct

Fraction	Volume	dilution	39	40	62	63
			type 1b	type 1b	type 3a	type 5a
20	250µl	1/200	0.072	0.130	0.096	0.051
21			0.109	0.293	0.084	0.052
22			0.279	0.249	0.172	0.052
23			0.093	0.151	0.297	0.054
24			0.080	0.266	0.438	0.056
25			0.251	0.100	0.457	0.048
26			3	1.649	0.722	0.066
27			3	3	2.528	0.889
28			3	3	3	2.345
29			3	3	2.849	2.580
30			2.227	1.921	1.424	1.333
31			0.263	0.415	0.356	0.162
32			0.071	0.172	0.154	0.064
33			0.103	0.054	0.096	0.057
34			0.045	0.045	0.044	0.051
35			0.043	0.047	0.045	0.046
36			0.045	0.045	0.049	0.040
37			0.045	0.047	0.046	0.048
38			0.046	0.048	0.047	0.057
39			0.045	0.048	0.050	0.057
40			0.046	0.049	0.048	0.049

Figure 24

34 / 69

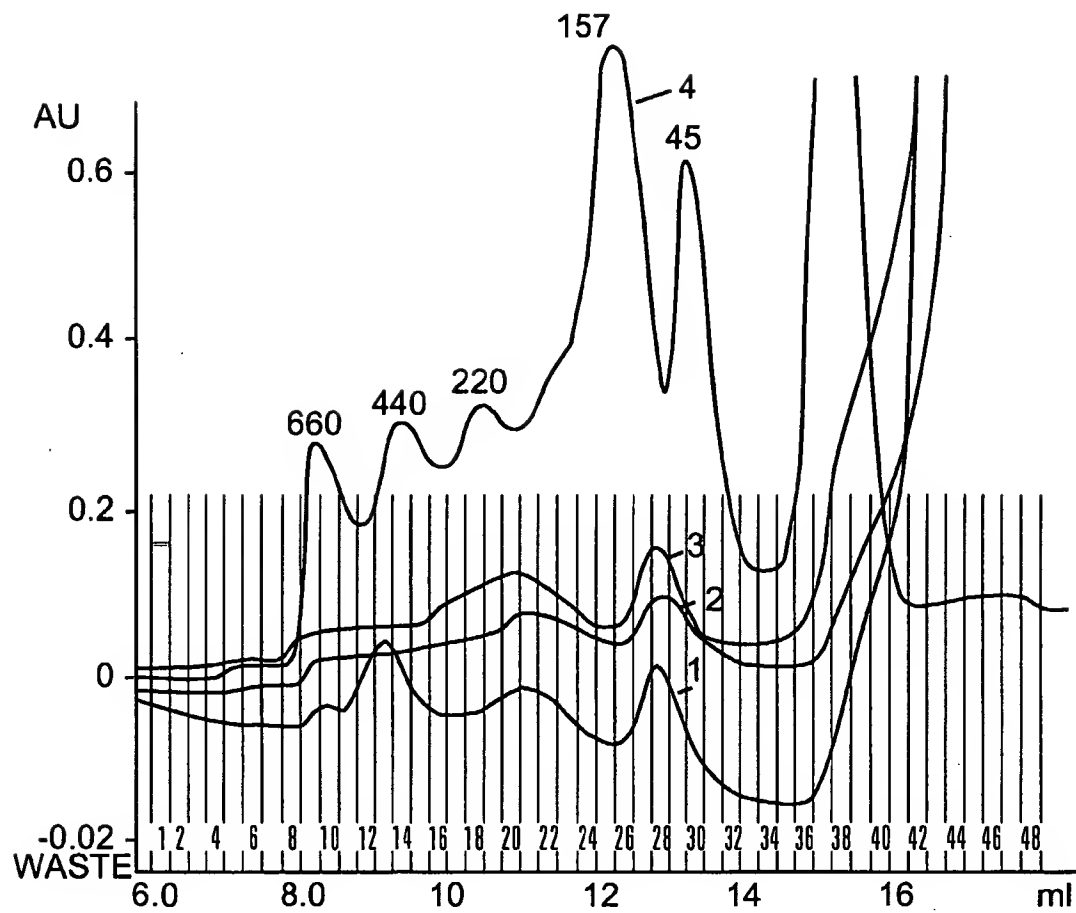


Figure 25

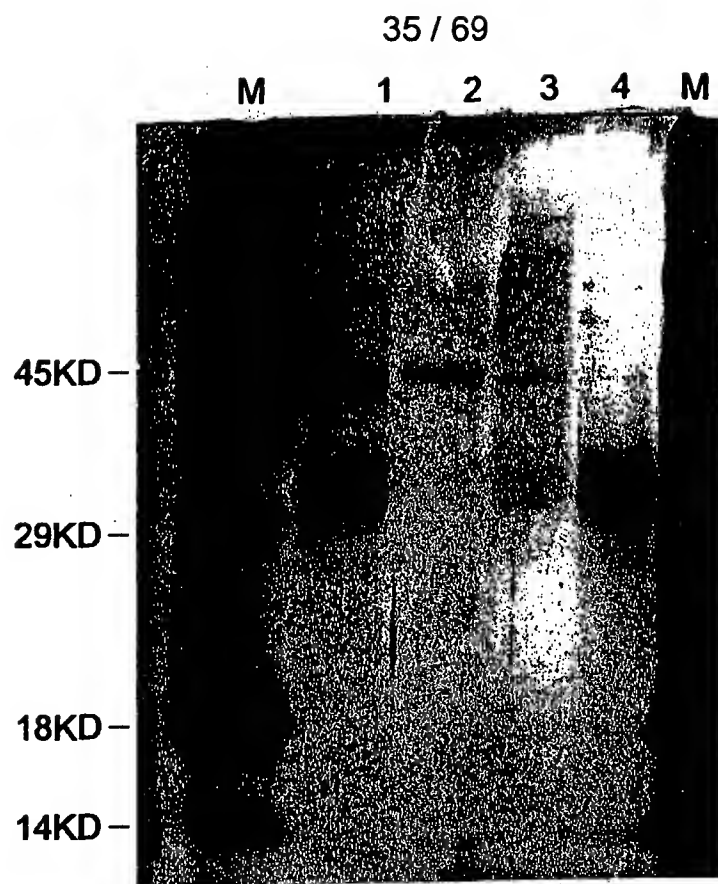


Figure 26

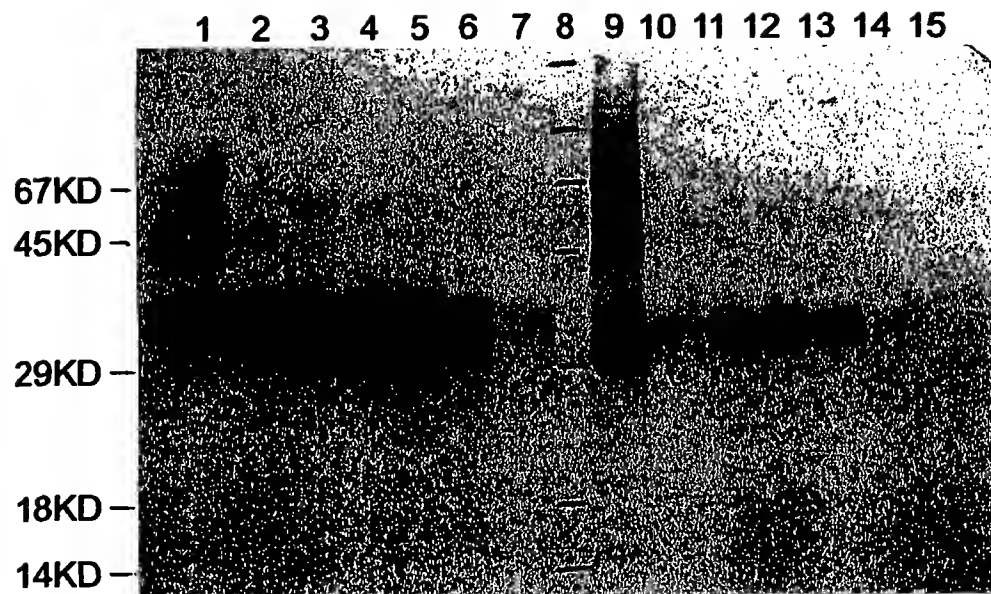
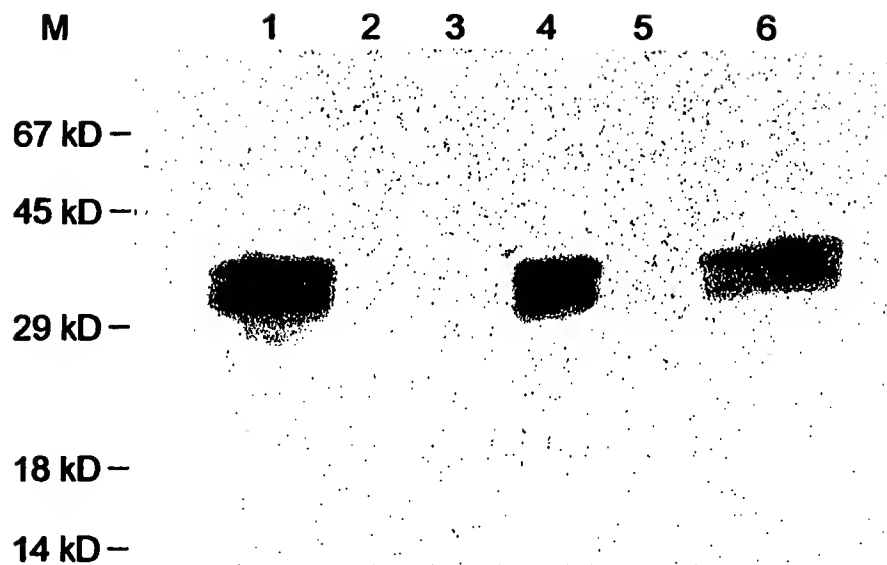


Figure 27

36 / 69



Figure 28



Lane 1: Crude Lysate  
Lane 2: Flow through Lentil Chromatography  
Lane 3: Wash with EMPIGEN Lentil Chromatography  
Lane 4: Eluate Lentil Chromatography  
Lane 5: Flow through during concentration lentil eluate  
Lane 6: Pool of E1 after Size Exclusion Chromatography

Figure 29

37 / 69

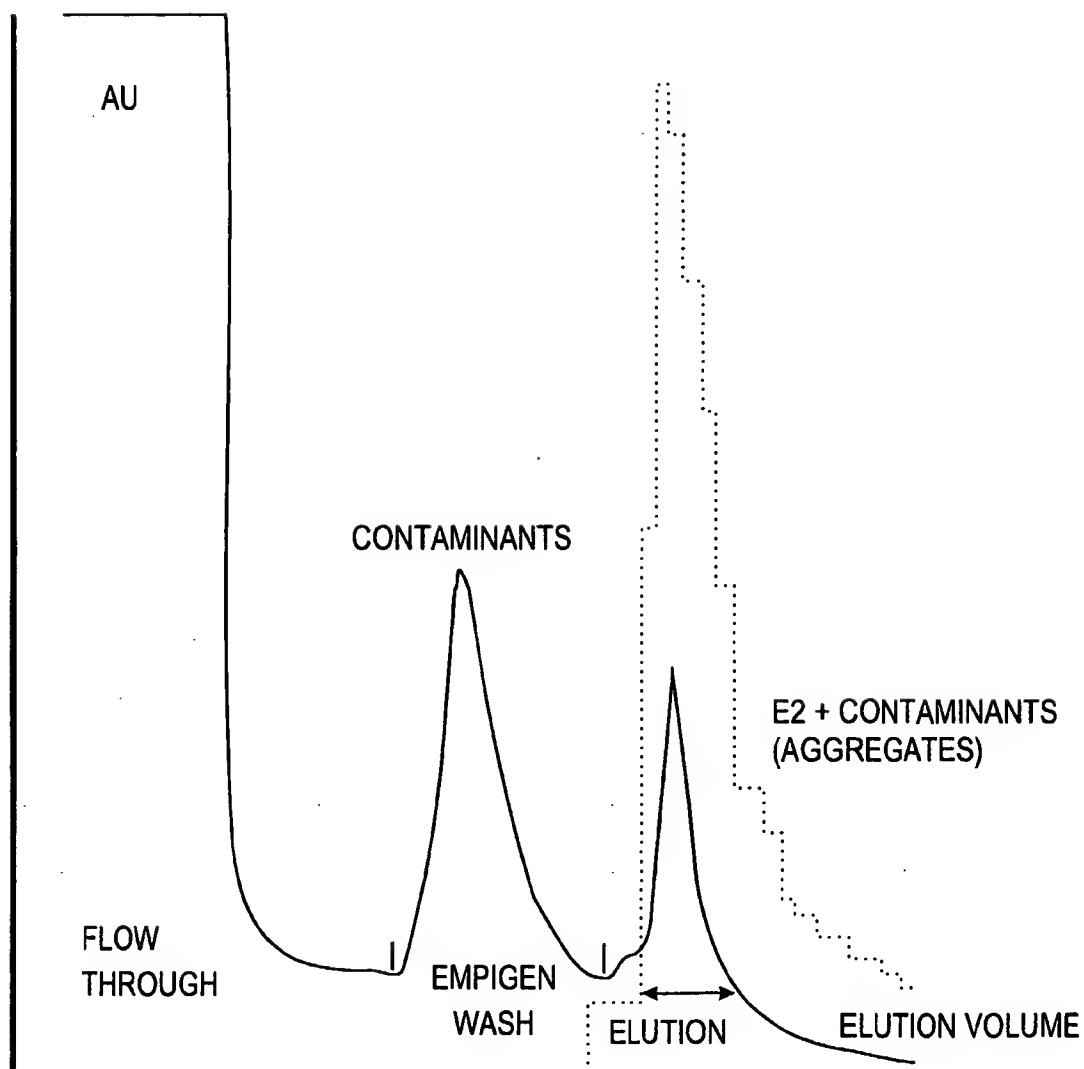


Figure 30

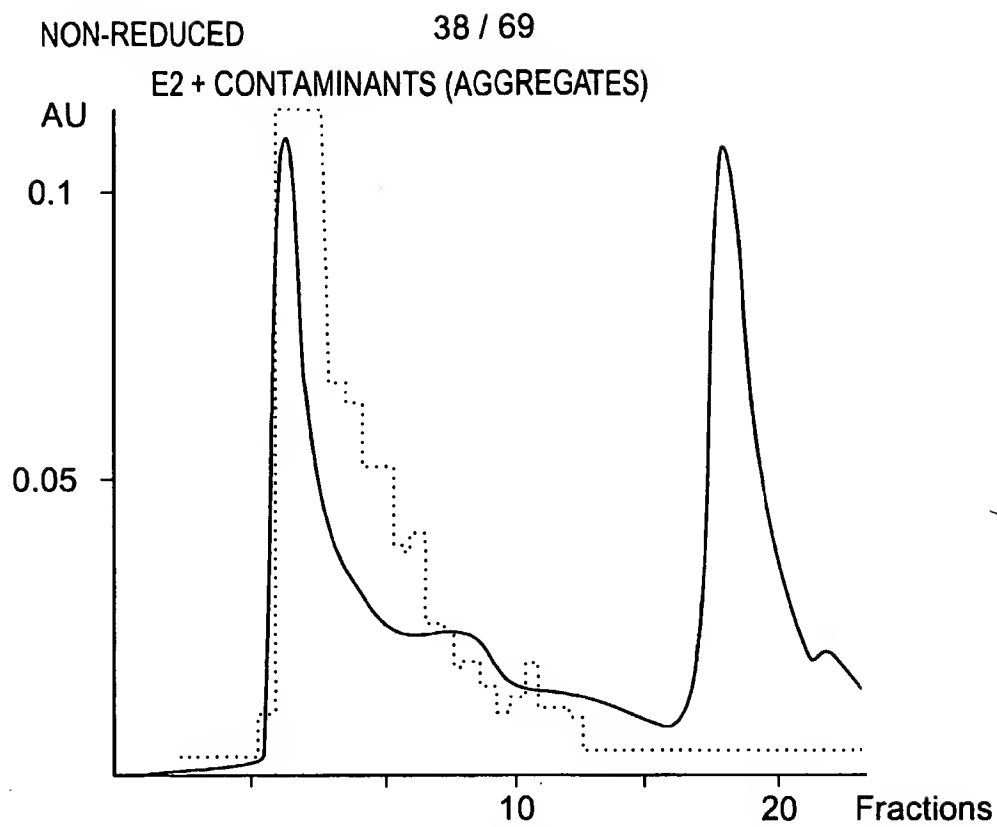


Figure 31A

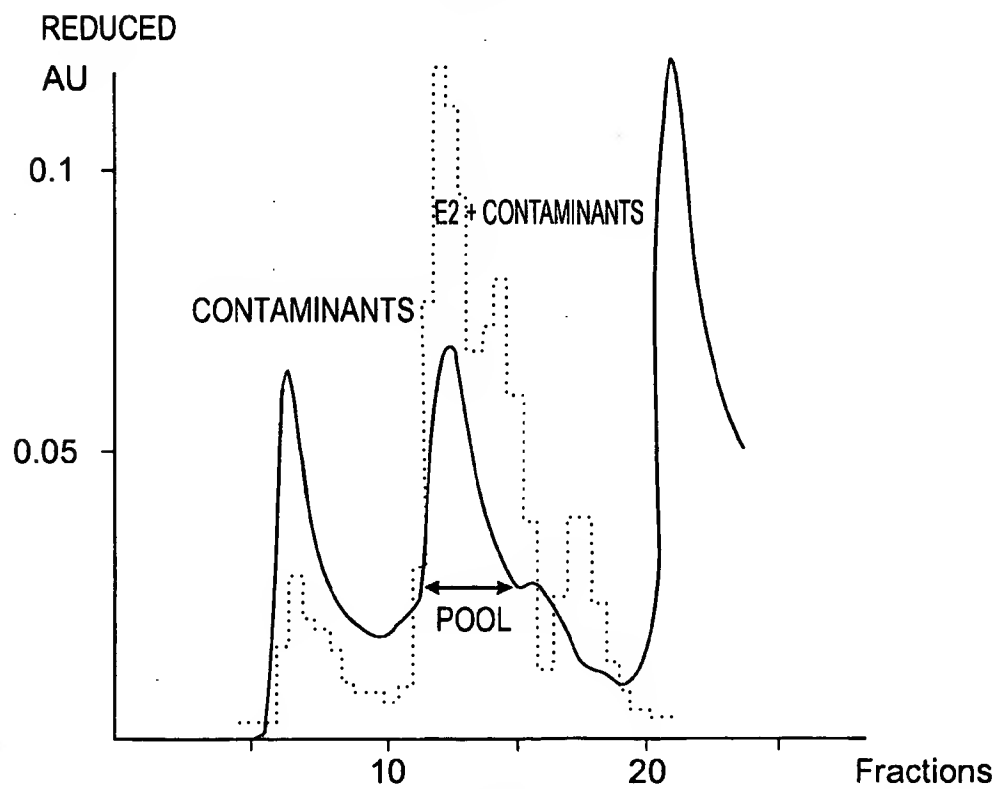


Figure 31B

39 / 69

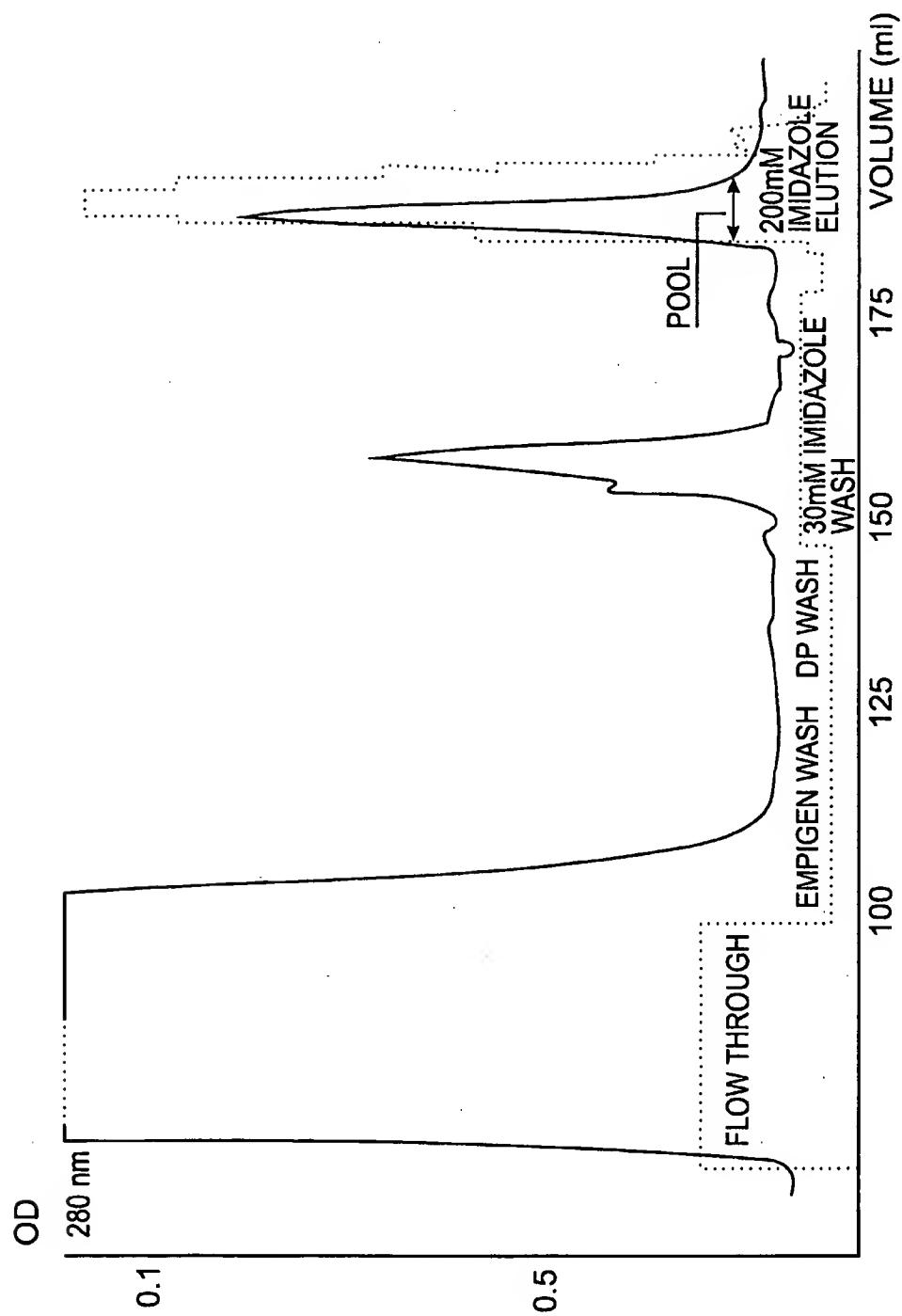
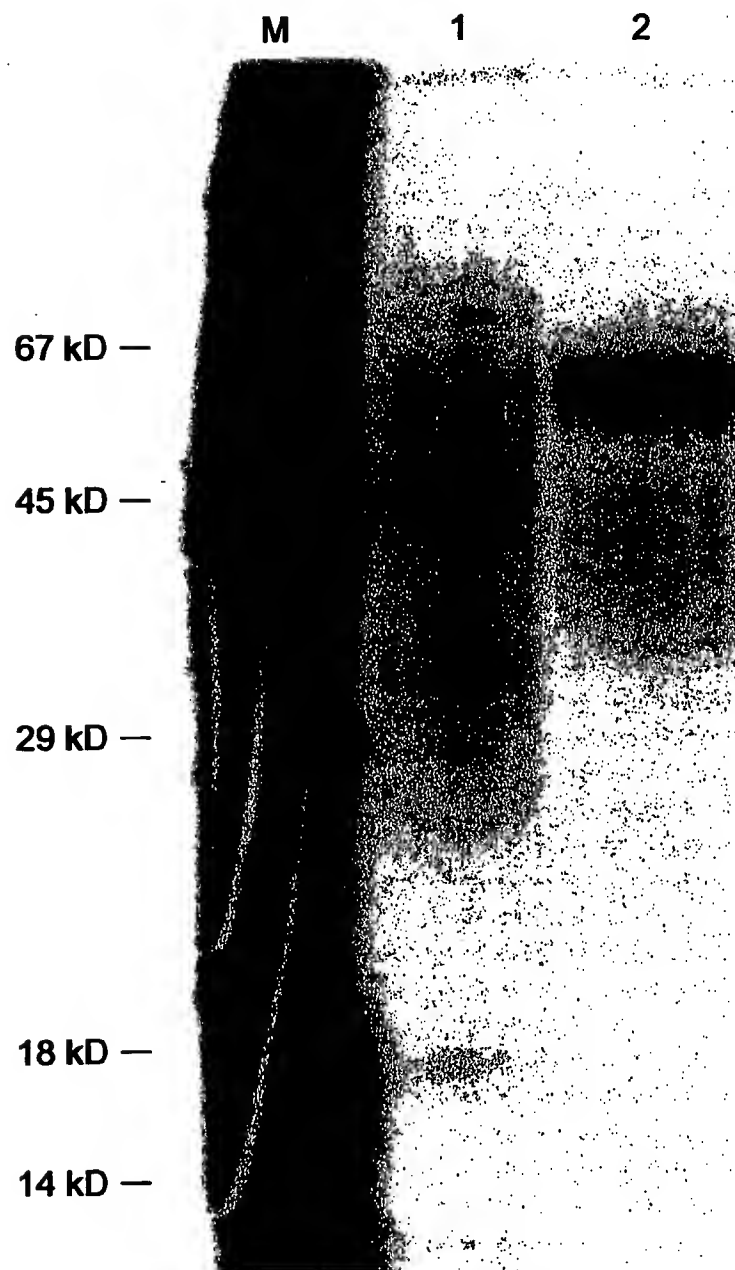


Figure 32

40 / 69

Figure 33: Silver stain of purified E2

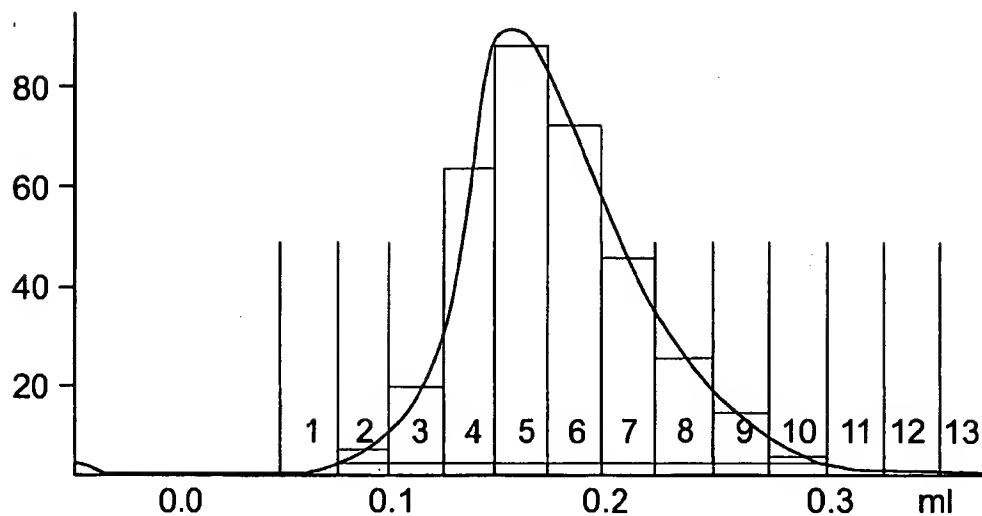


1. 30mM imidazole wash Ni-IMAC  
2. 0.5µg E2

Figure 33



41 / 69



No.	Ret (ml)	Peak start (ml)	Peak end (ml)	Dur (ml)	Area (ml*mAU)	Height (mAU)
1	-0.45	-0.46	-0.43	0.04	0.0976	4.579
2	1.55	0.75	3.26	2.51	796.4167	889.377
3	3.27	3.26	3.31	0.05	0.0067	0.224
4	3.33	3.32	3.33	0.02	0.0002	0.018

Total number of detected peaks = 4

Total Area above baseline = 0.796522 ml\*AU

Total area in evaluated peaks = 0.796521 ml\*AU

Ratio peak area / total area = 0.999999

Total peak duration = 2.613583 ml

Figure 34

42 / 69

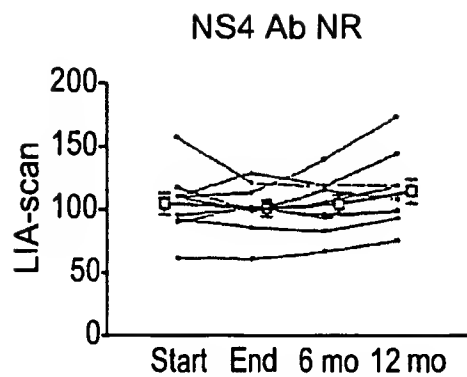


Fig. 35A-1

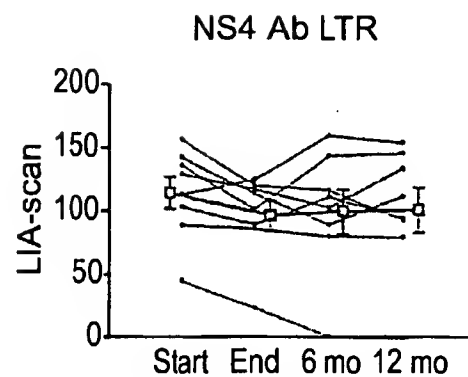


Fig. 35A-2

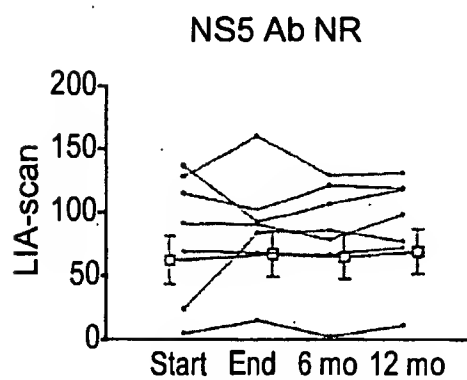


Fig. 35A-3

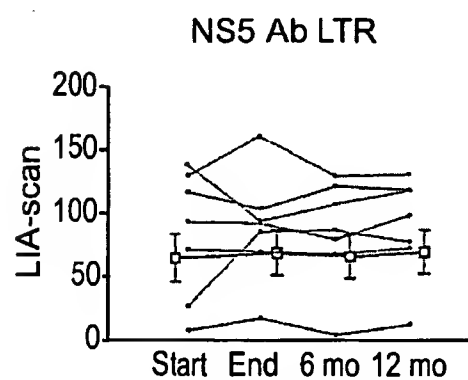


Fig. 35A-4

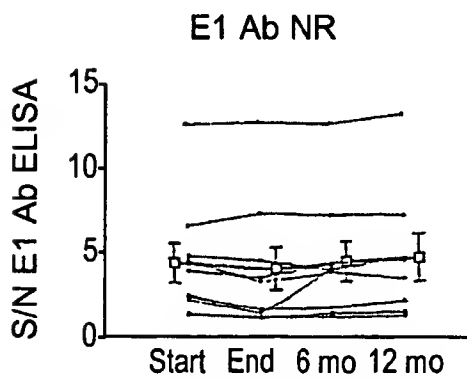


Fig. 35A-5

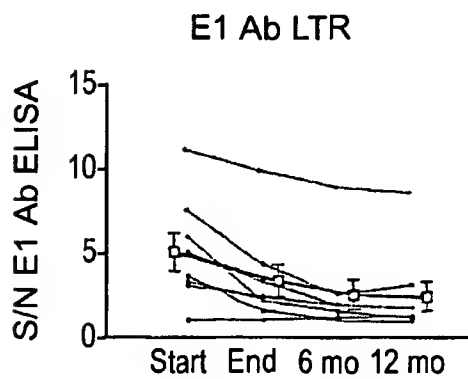


Fig. 35A-6

43 / 69

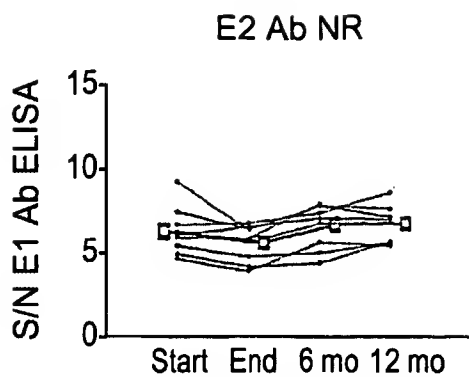


Fig. 35A-7

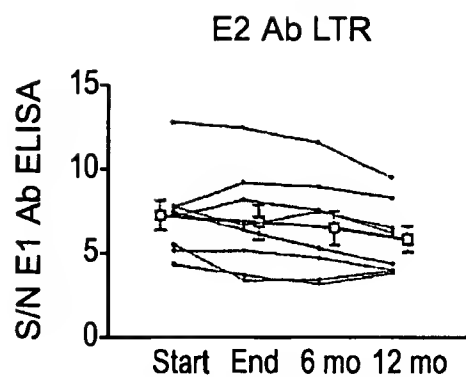


Fig. 35A-8

44 / 69

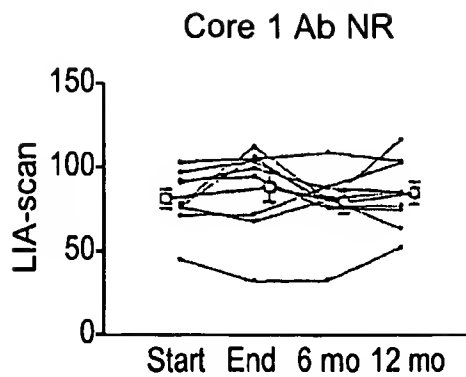


Fig. 35B-1

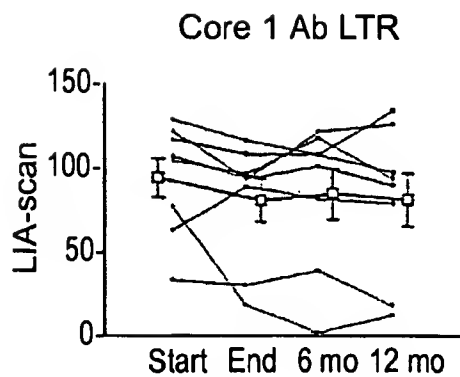


Fig. 35B-2

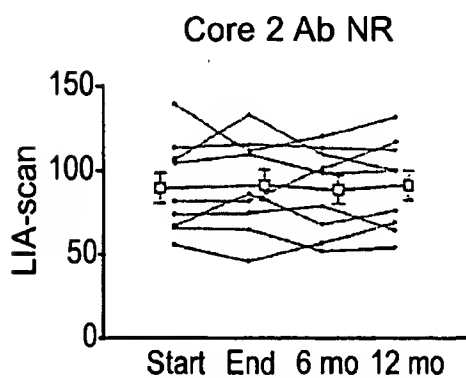


Fig. 35B-3

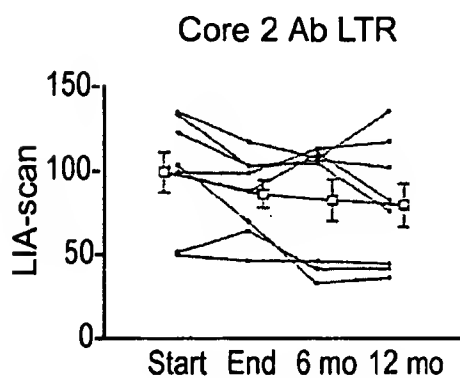


Fig. 35B-4

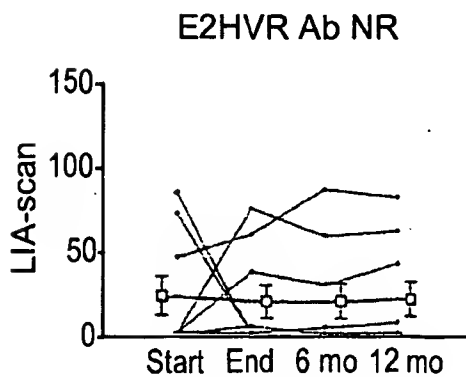


Fig. 35B-5

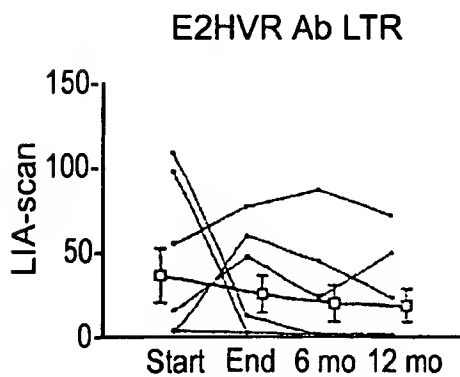


Fig. 35B-6

45 / 69

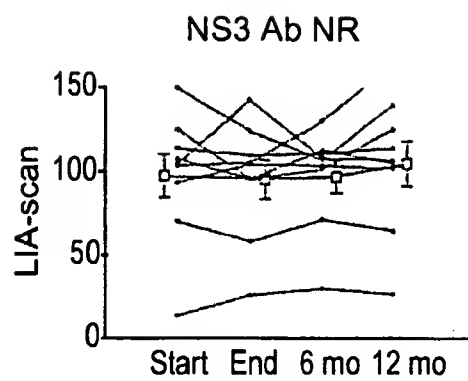


Fig. 35B-7

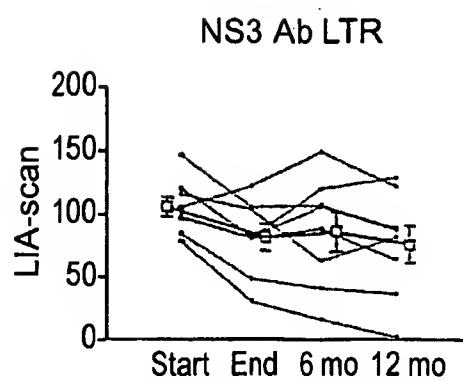


Fig. 35B-8

46 / 69

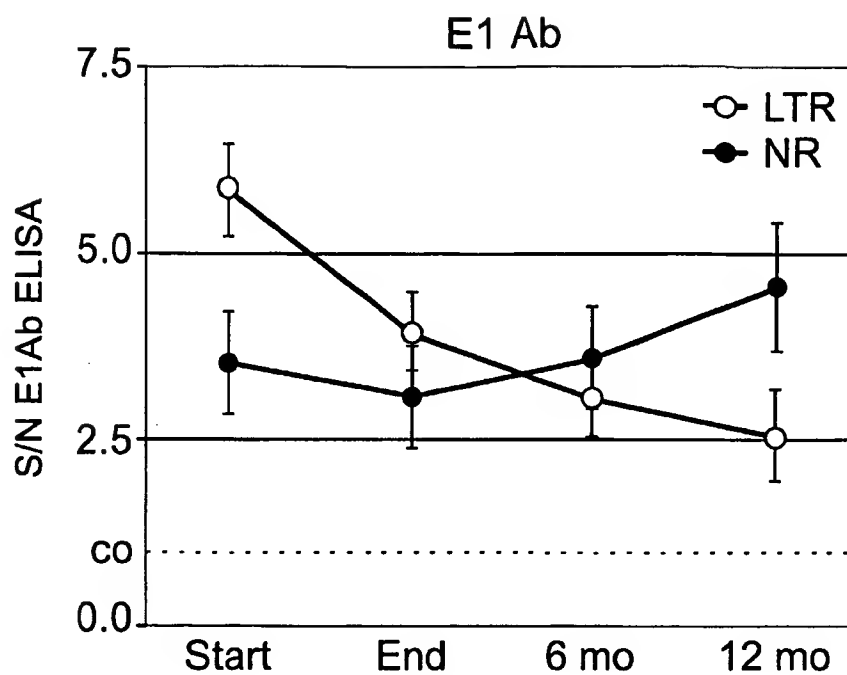


Figure 36A

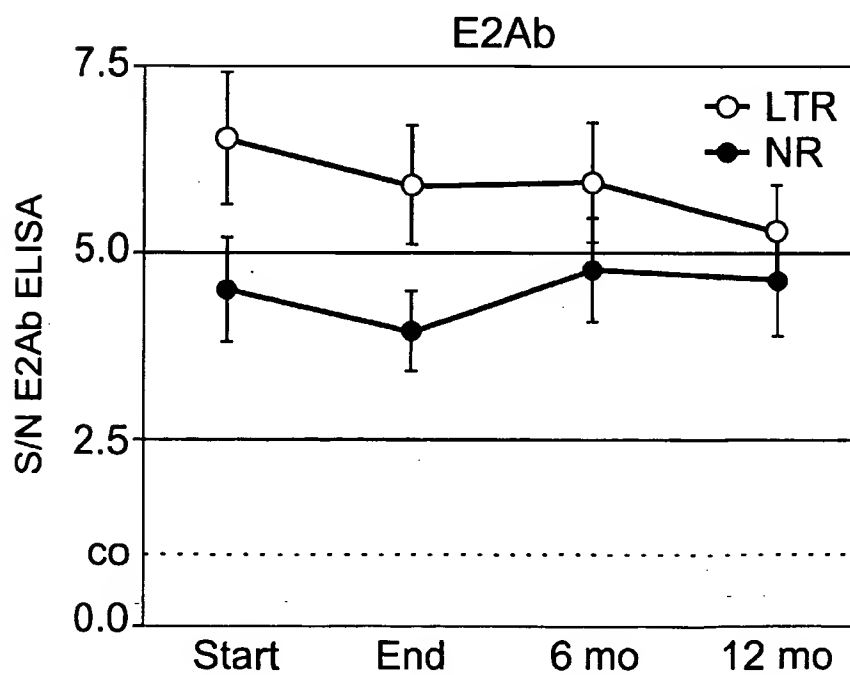


Figure 36B

47 / 69

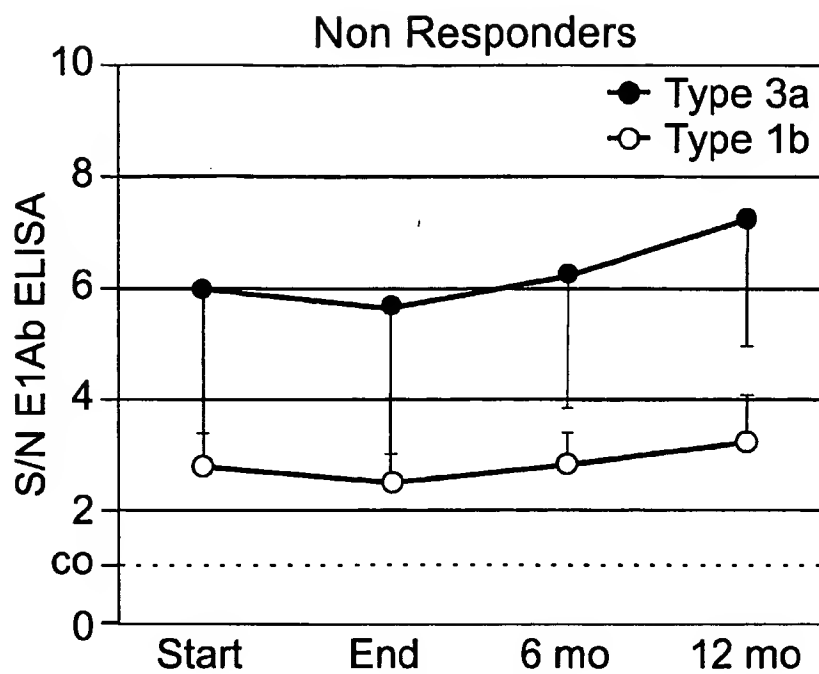


Figure 37A

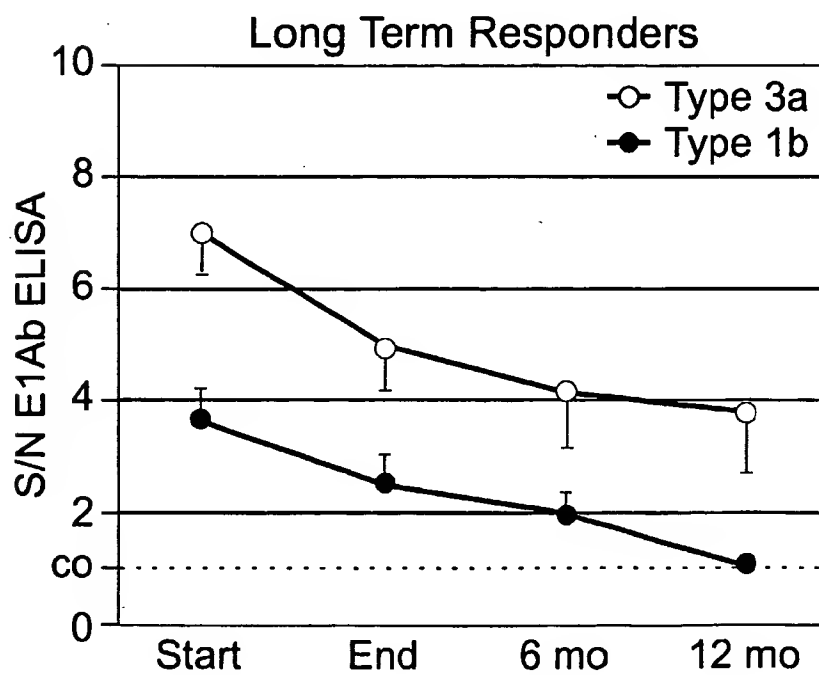


Figure 37B

48 / 69

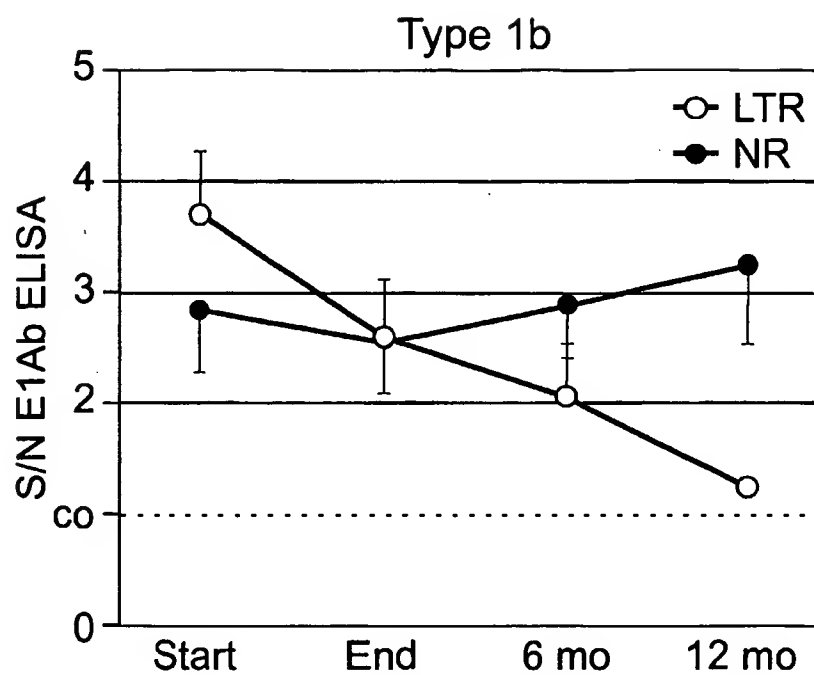


Figure 37C

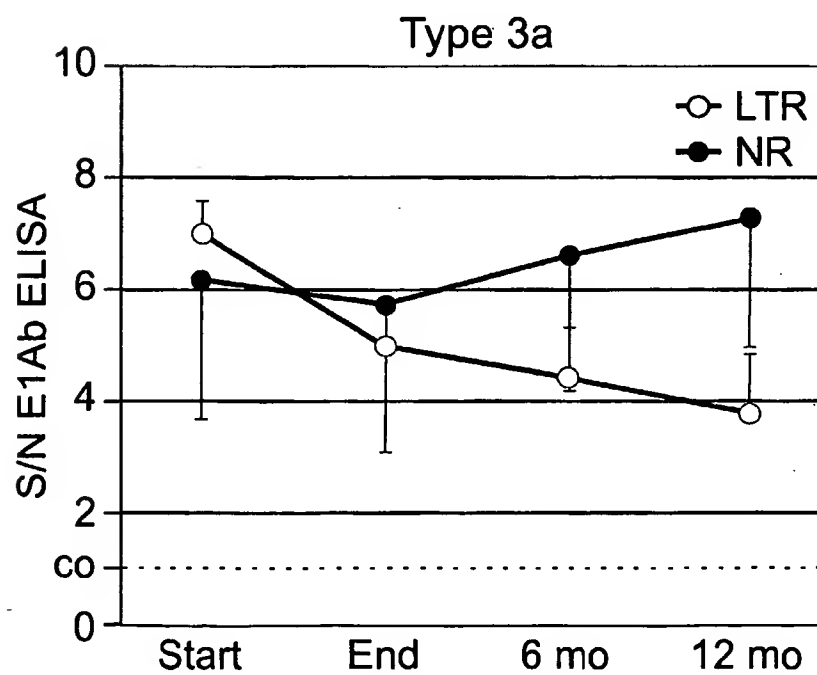


Figure 37D



49 / 69

Relative Map Positions of  
anti-E2 monoclonal antibodies

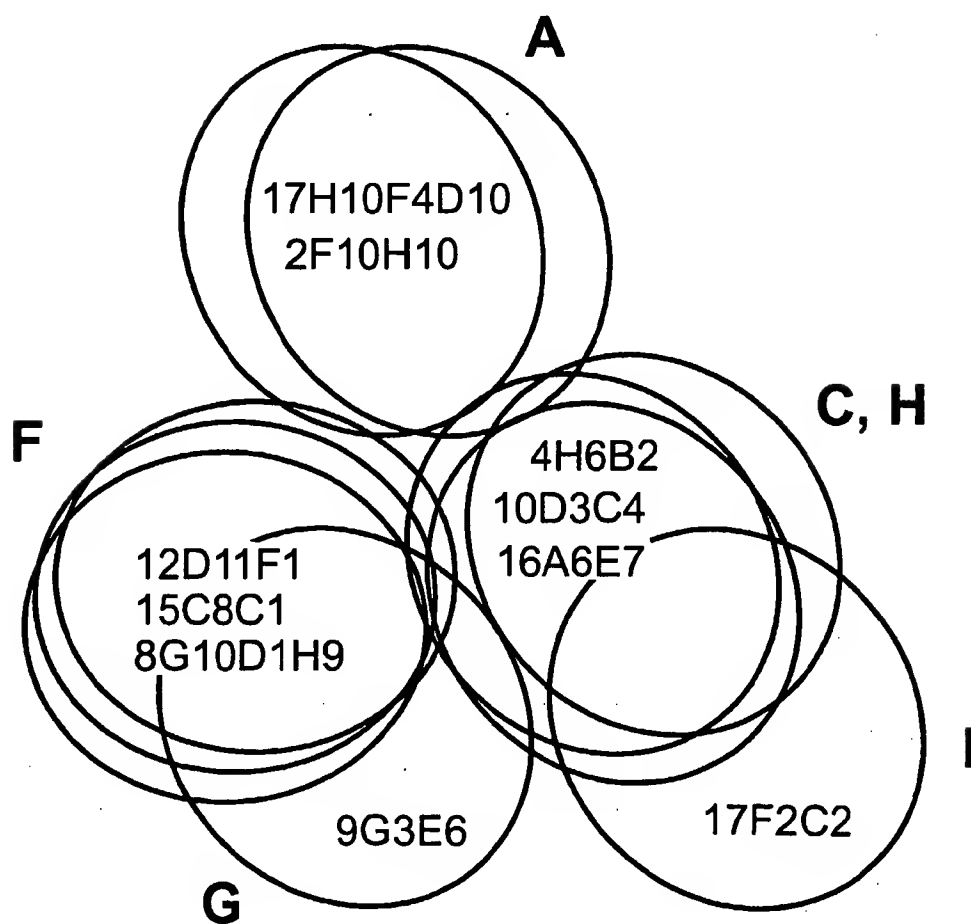


Figure 38

50 / 69

# PARTIAL DEGLYCOSYLATION OF HCV E1 ENVELOPE PROTEIN

Endoglycosidase H  
(Endo H)

Glycopeptidase F  
(PNGase F)

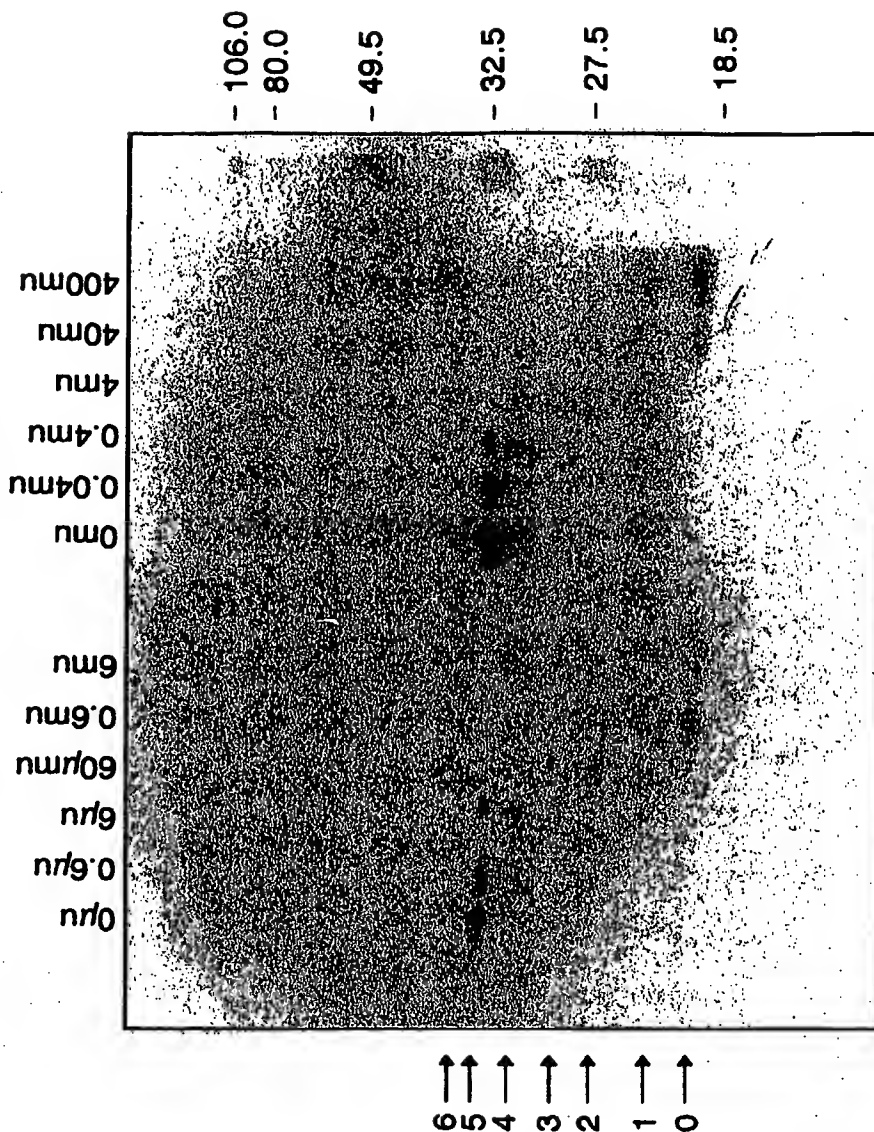


Figure 39

51 / 69

# **PARTIAL TREATMENT OF HCV E2\E2s ENVELOPE PROTEINS BY PNGase F**

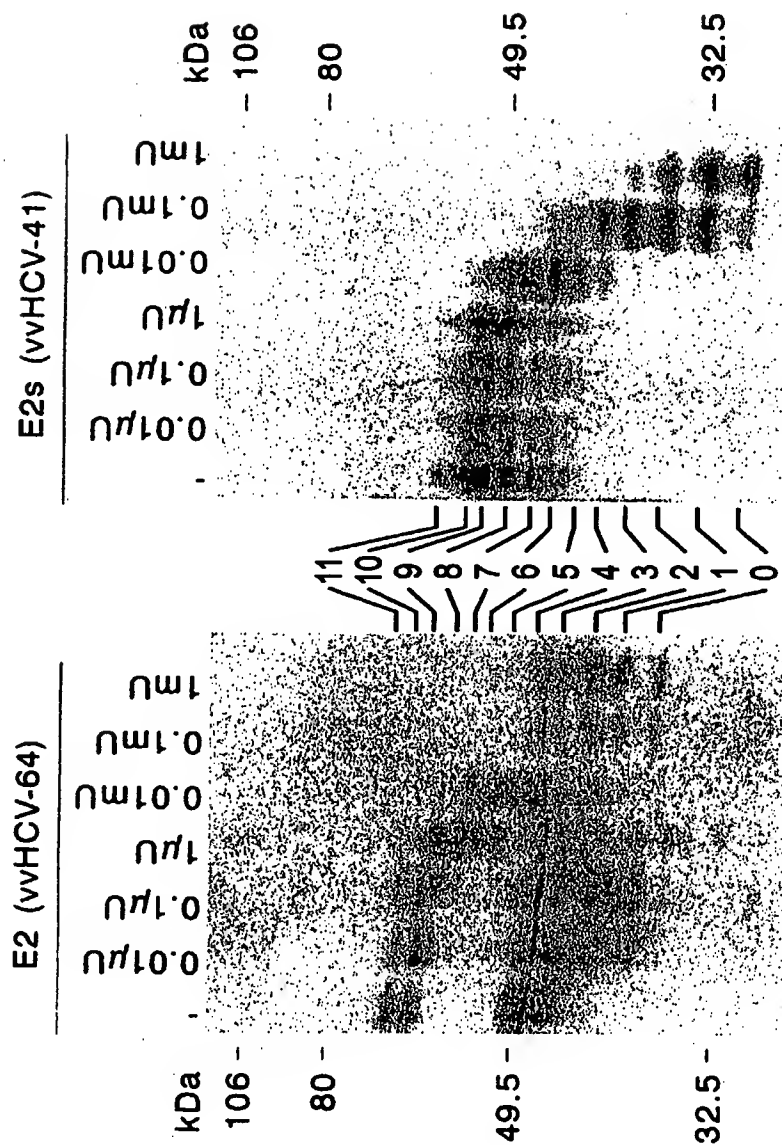


Figure 40

In vitro mutagenesis of HCV E1 glycoprotein

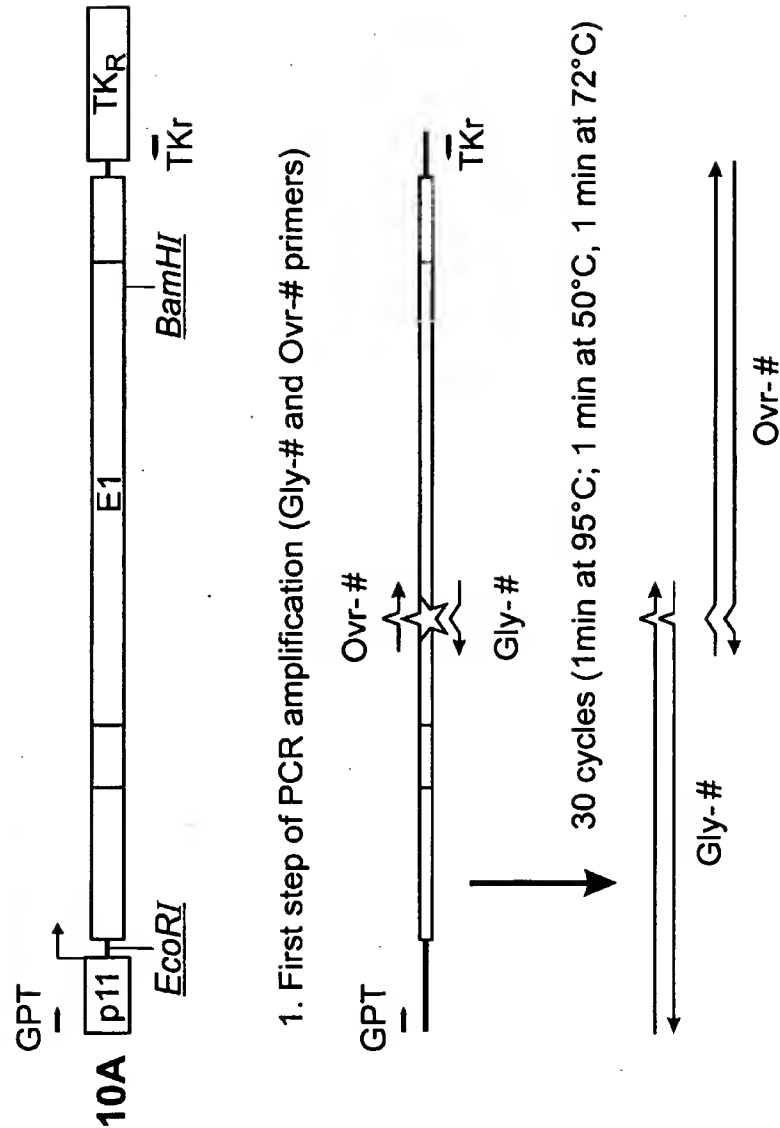


Figure 42A

54 / 69

## 2. Overlap extension and nested PCR

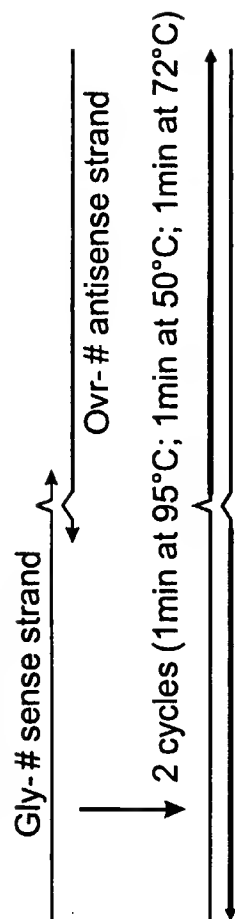
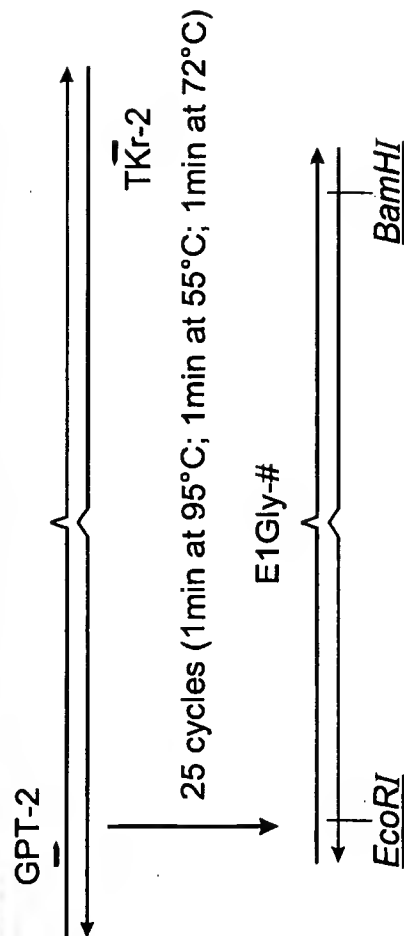
a. Overlap extensionb. Nested PCR amplification (GPT-2 and TKr-2 primers)

Figure 42B

## In vitro mutagenesis of HCV E1 glycoprotein

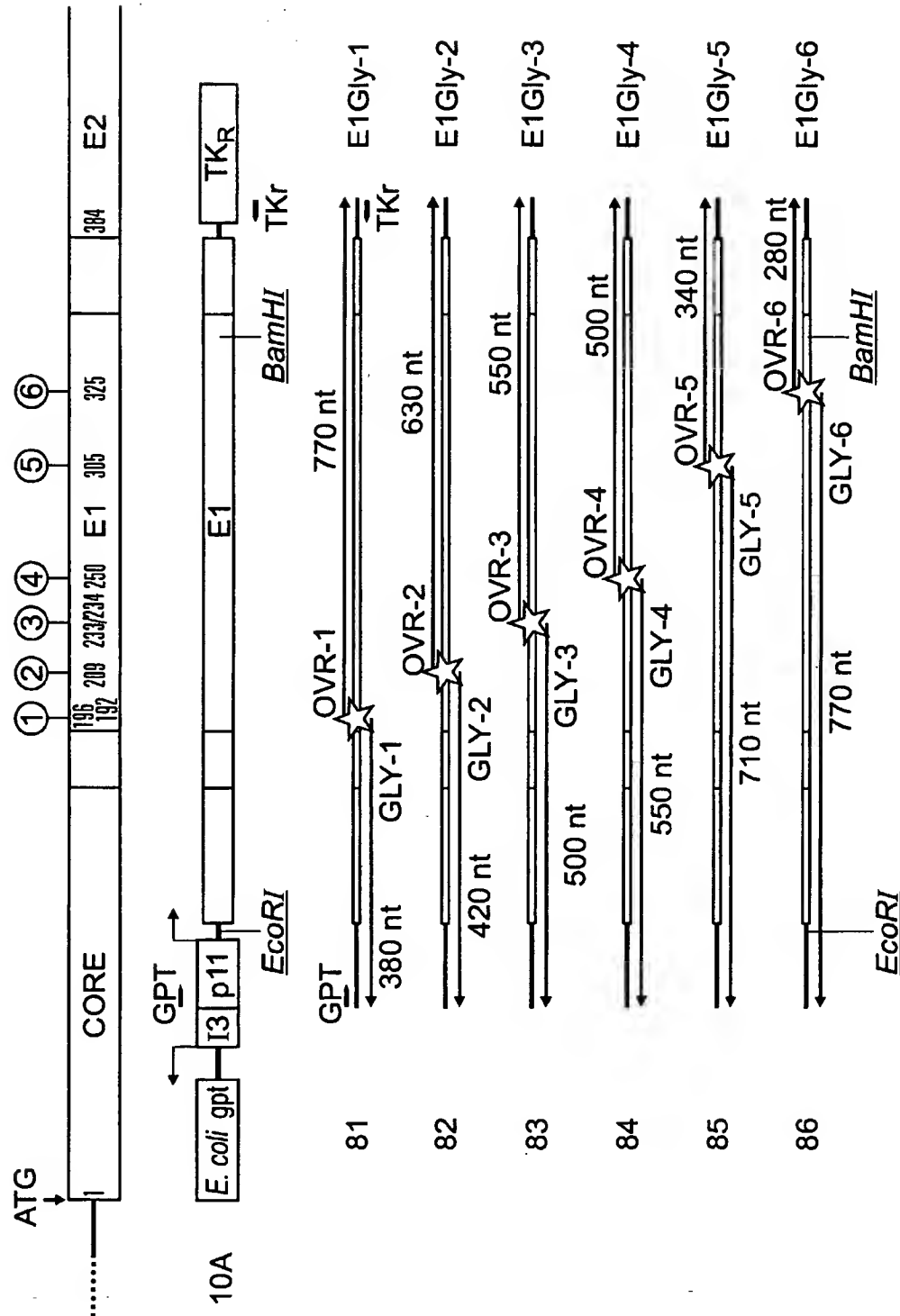


Figure 43

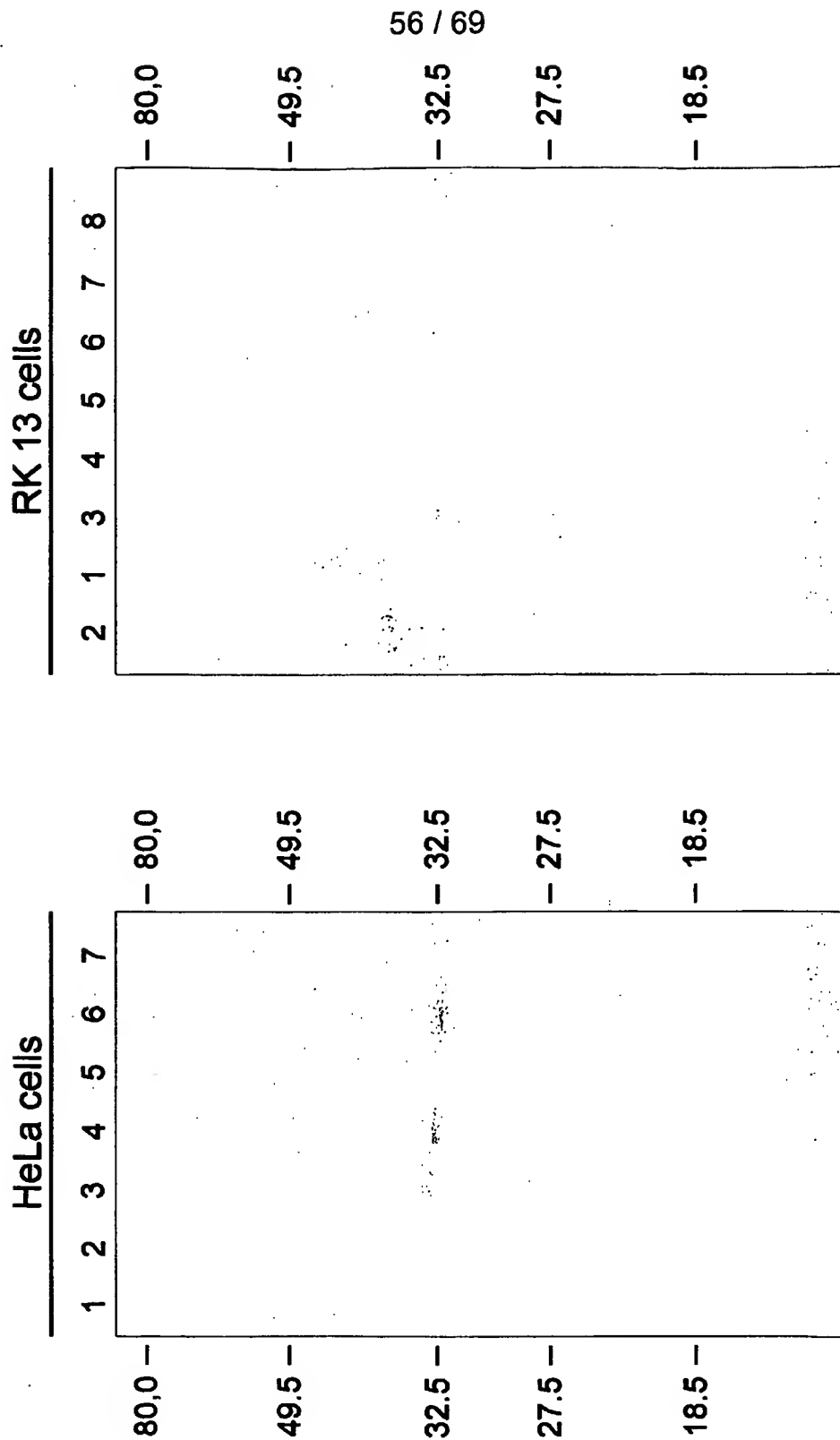


Figure 44A

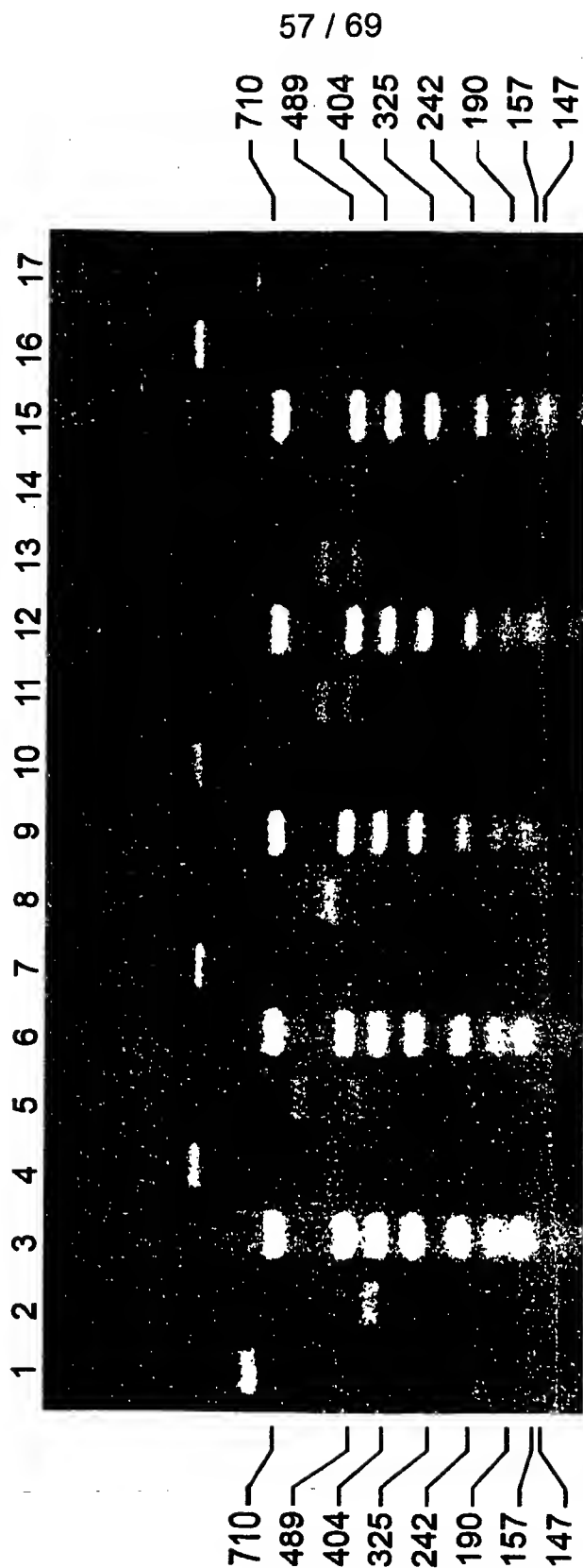


Figure 44B



58 / 69



Figure 45

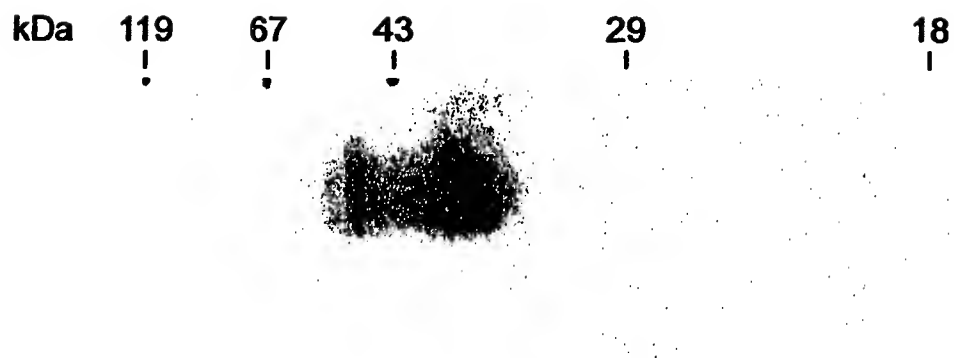


Figure 46

59 / 69

	age (years)	HCV infection (years)	genotype
Marcel	17	9	1a
Peggy	21	16,5	1b
Femma	15	9	1a
Yoran	12	none	
Marti	12	none	

**chronic carriers (strong T-cell adjuvant)**

↓ ↓ ↓ ↓ ↓ ↓      ↓ ↓ ↓ ↓ ↓ ↓      50 µg E1 dose  
 —————  
 0 3 6 9 12 15      26 29 32 35 38 41      weeks

**naive (alum)**

↓ ↓ ↓ ↓ ↓ ↓      50 µg E1 dose  
 —————  
 0 3 6 9 12 15      weeks

Figure 47

60 / 69

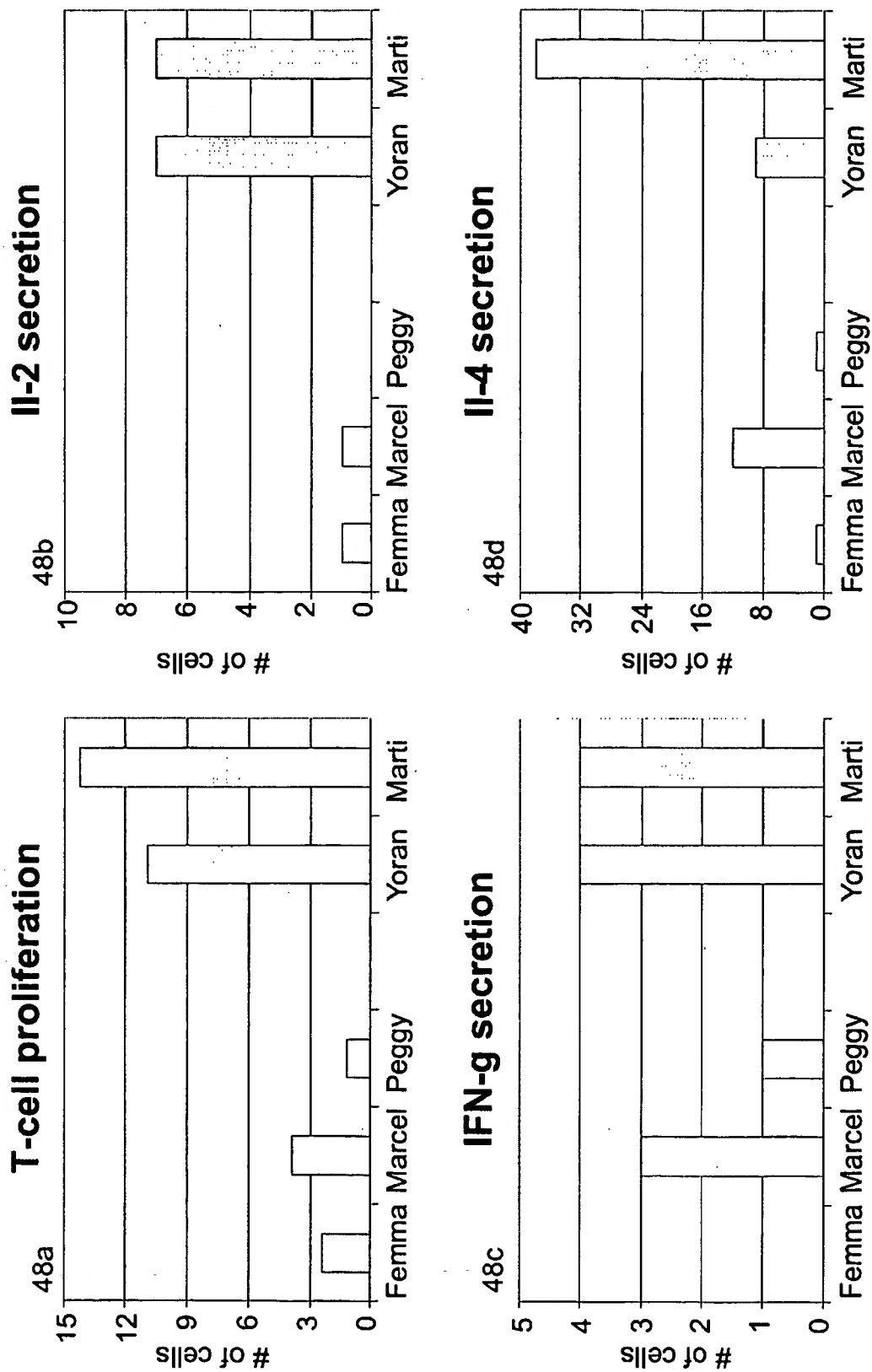


Figure 48

61 / 69

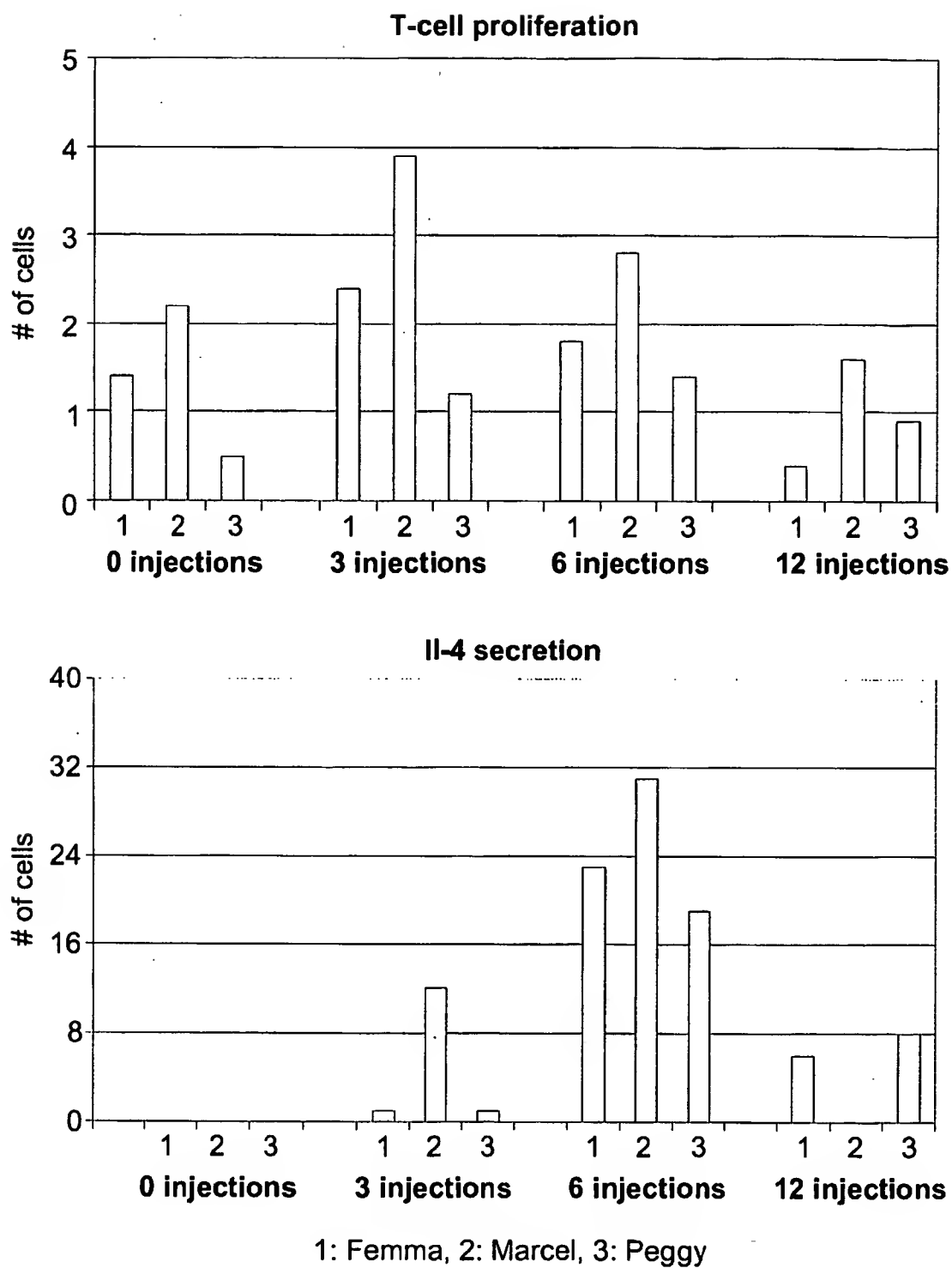


Figure 49

62 / 69

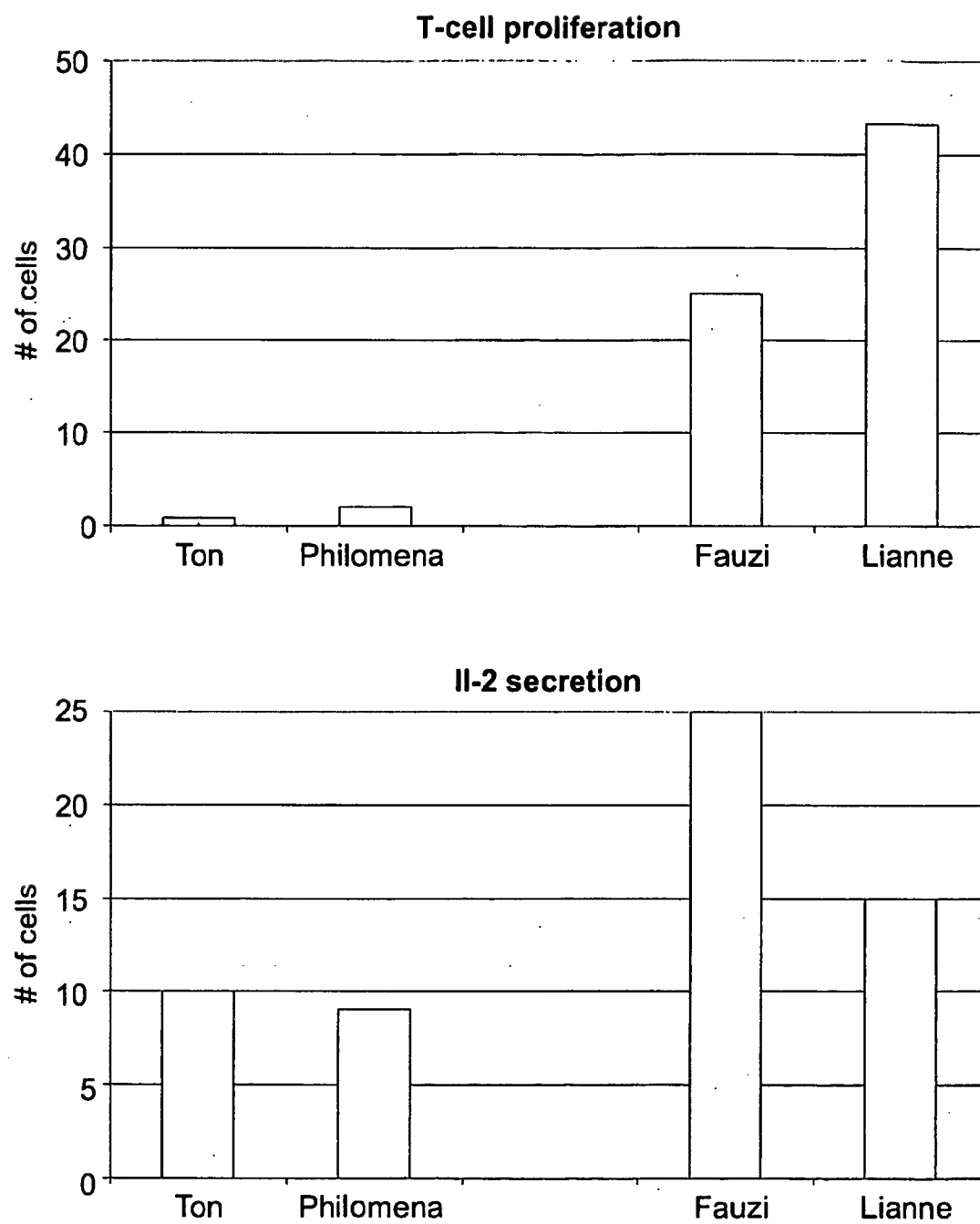


Figure 50

63 / 69

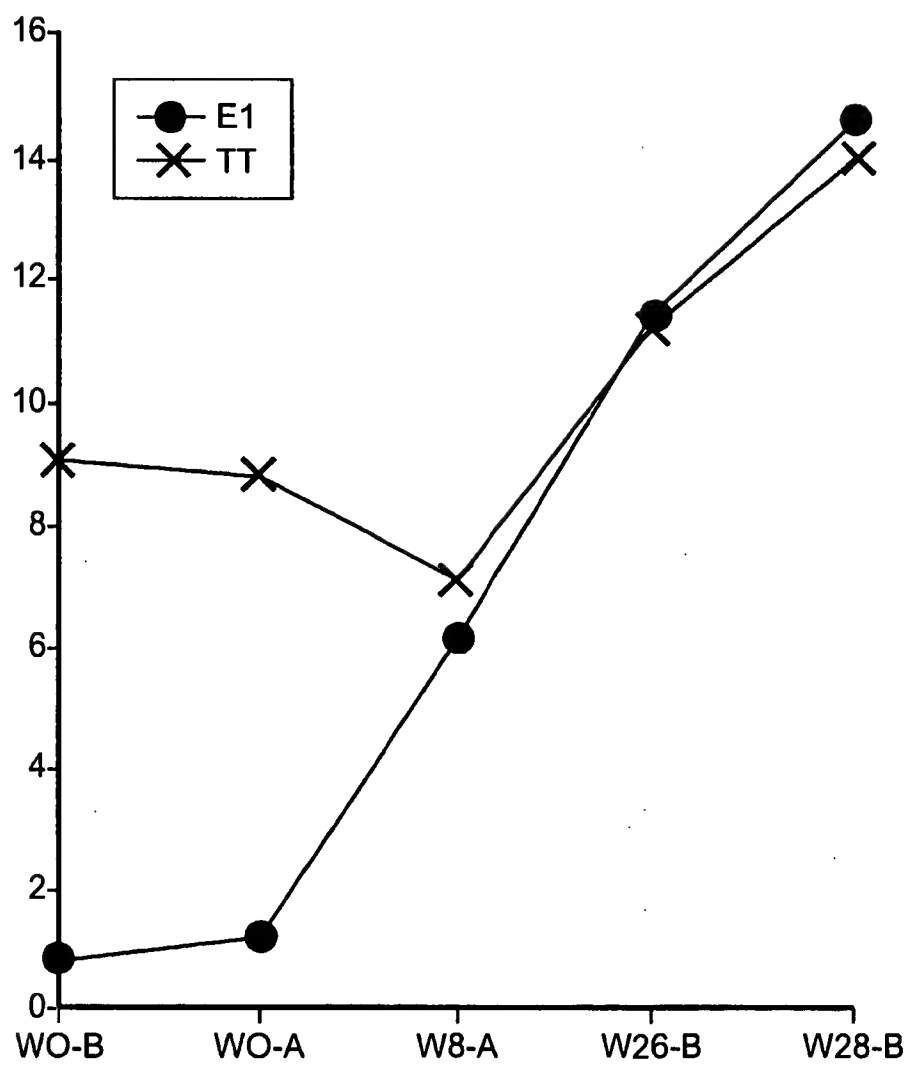


Figure 51

64 / 69

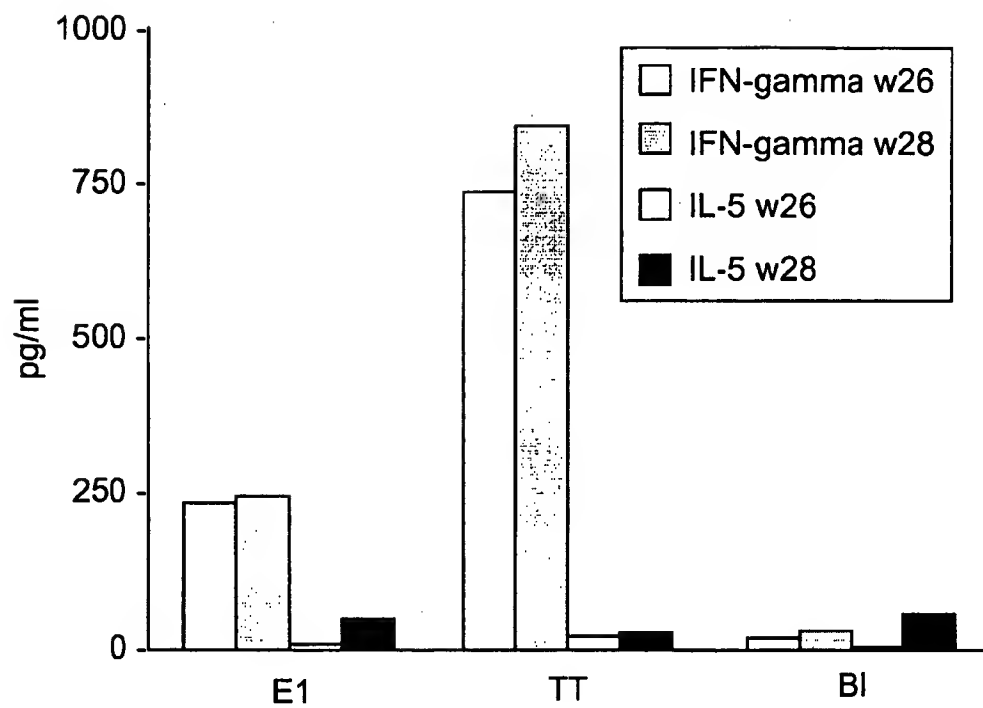


Figure 52

65 / 69

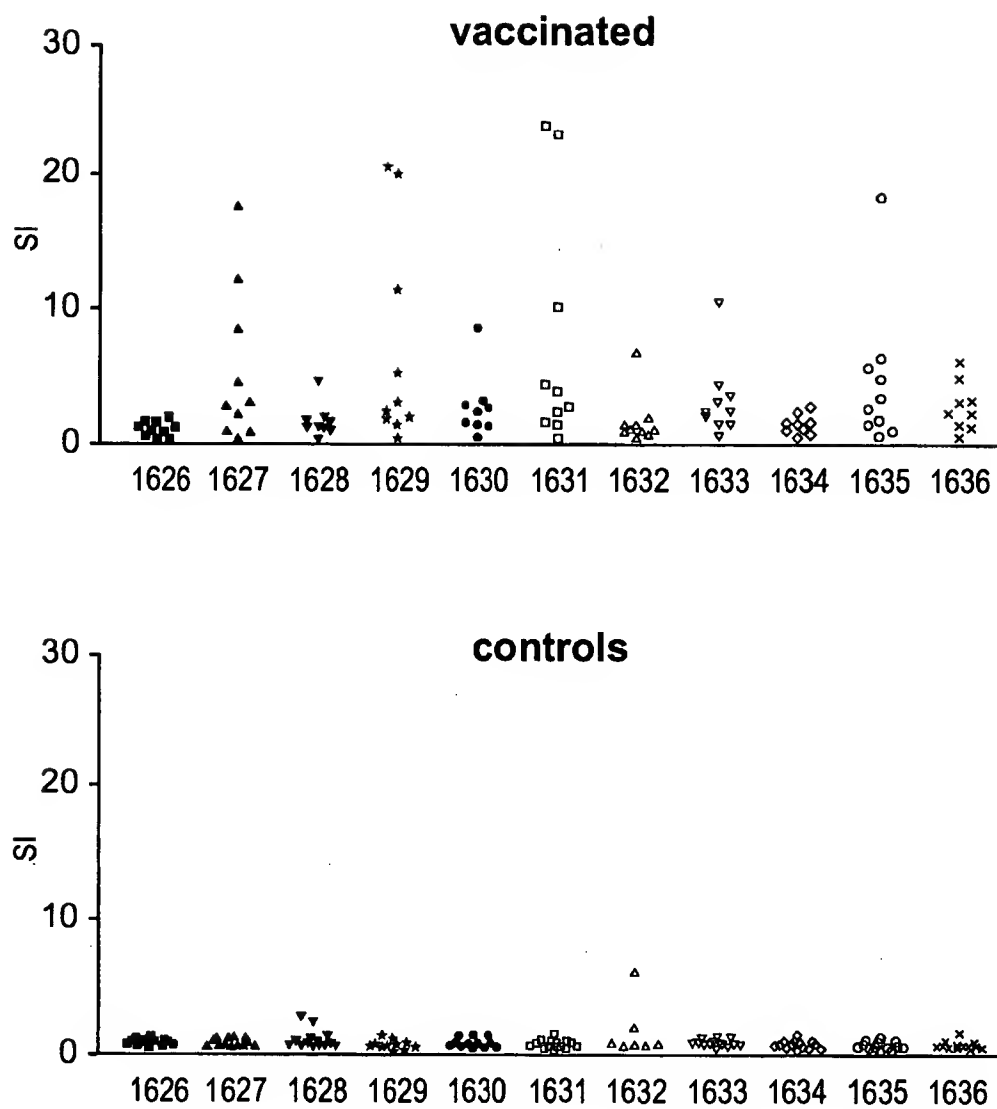


Figure 53



66/69

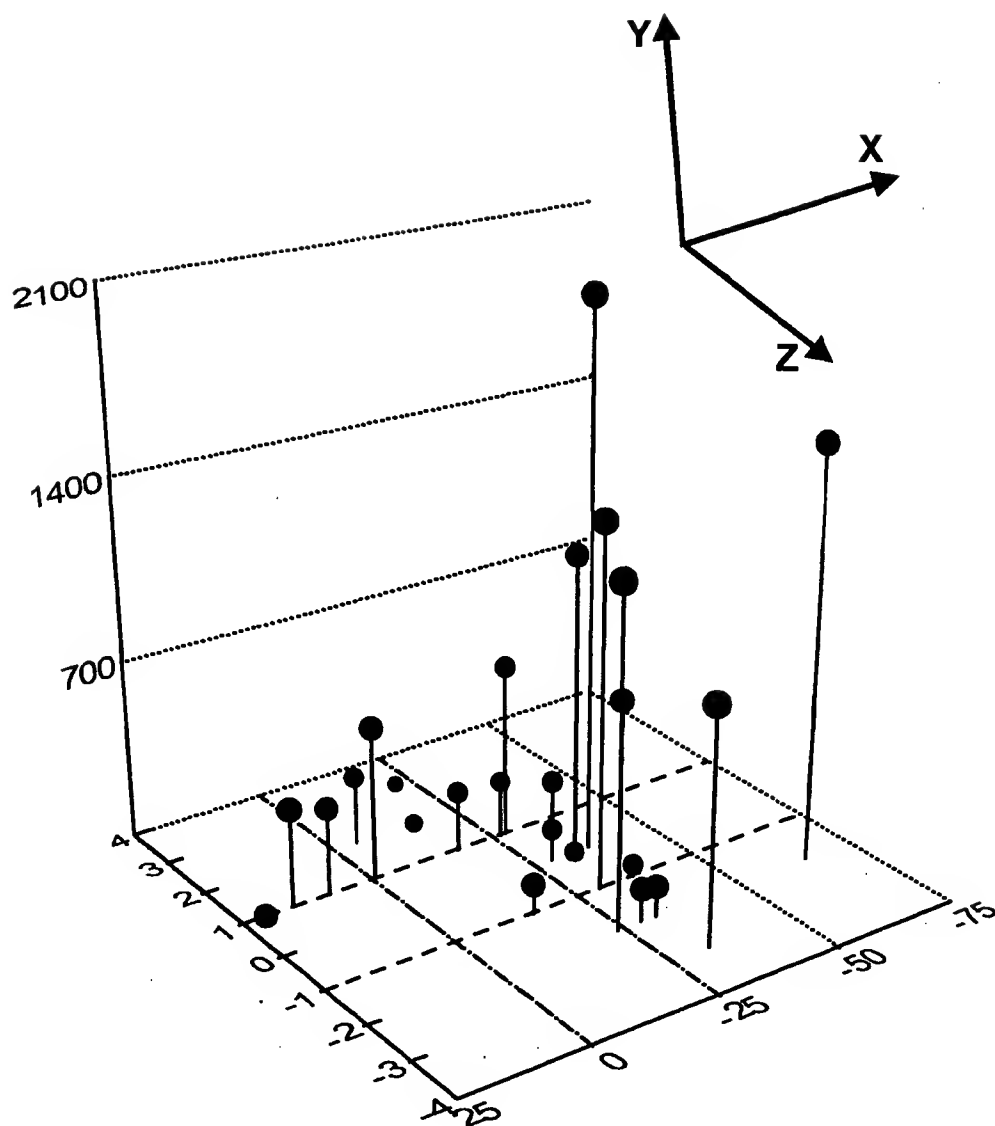


FIGURE 54

67/69

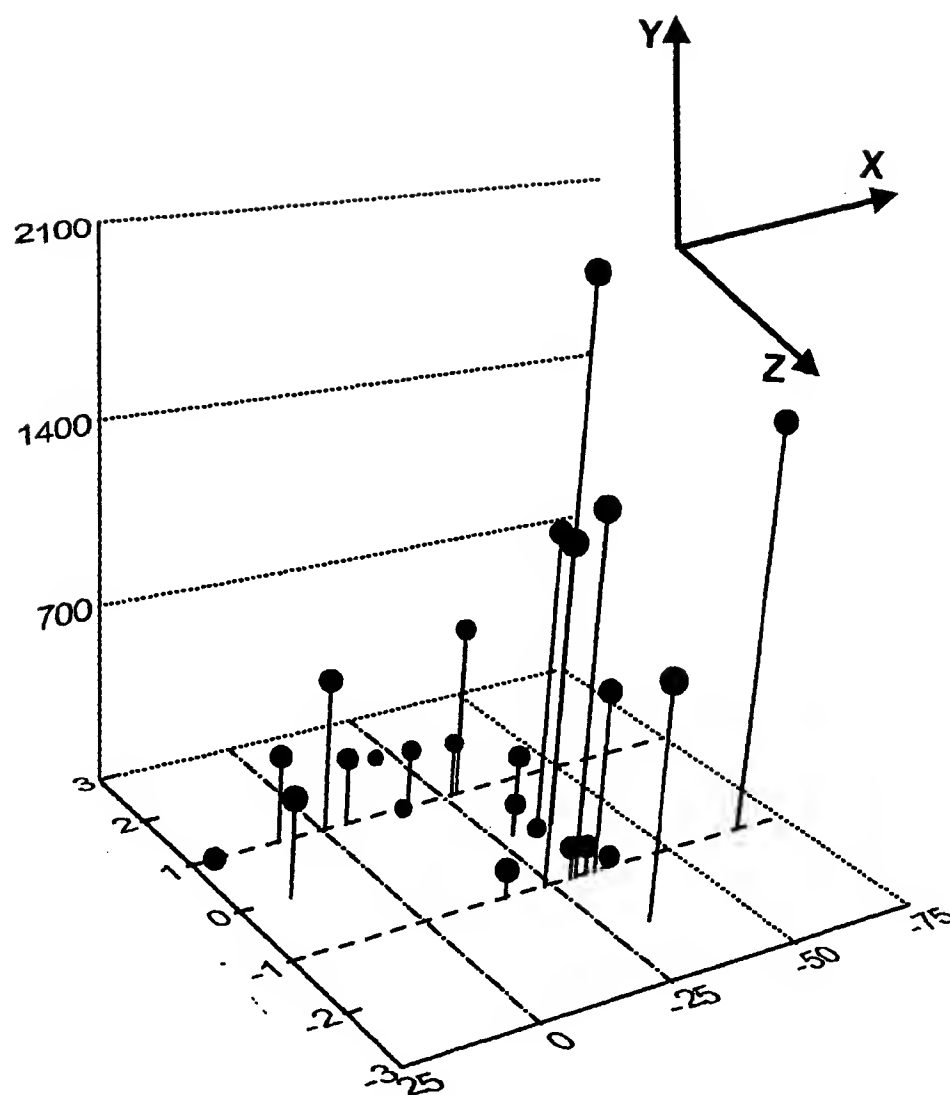


FIGURE 55

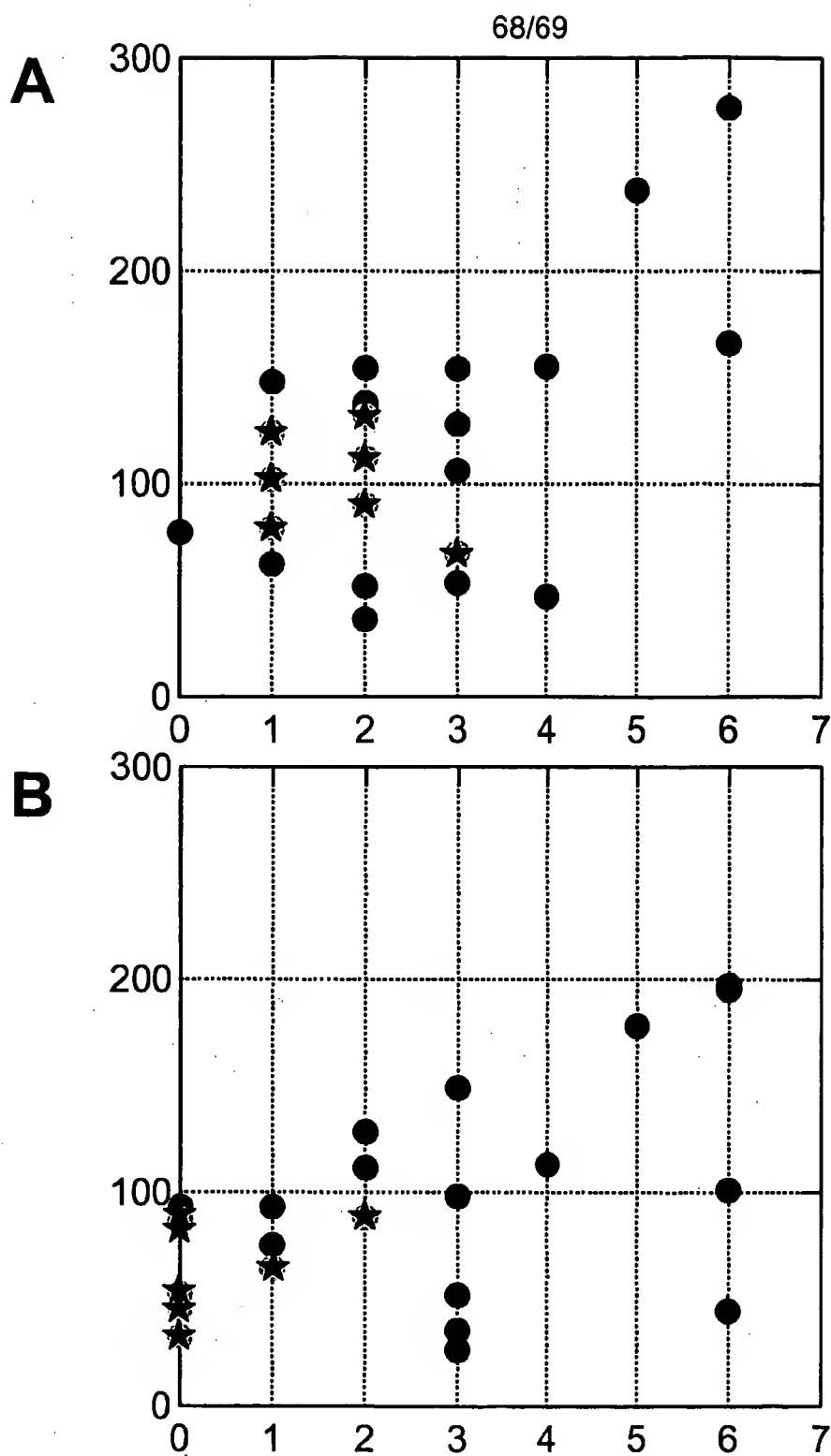


FIGURE 56

69/69

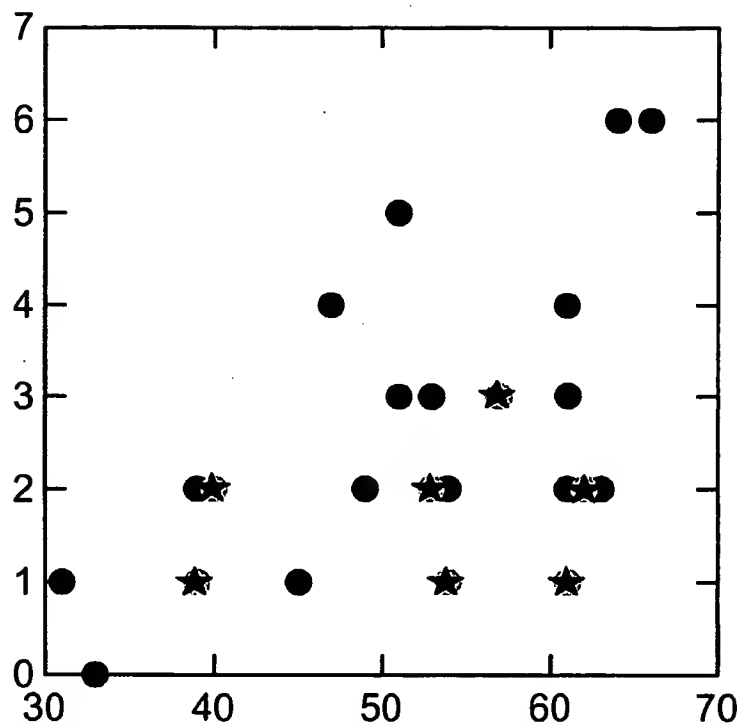
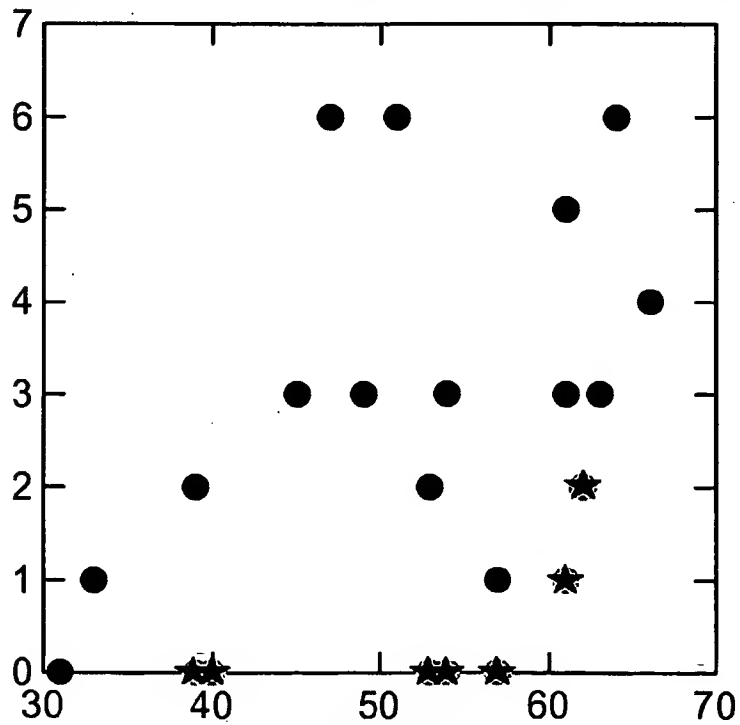
**A****B**

FIGURE 57

## Sequence Listing

<110> Innogenetics N.V.

<120> Purified hepatitis C virus envelope proteins for diagnostic and therapeutic use.

<140> PCT/EP02/00219

<160> 123

<170> PatentIn 3.1

<210> 1

<211> 21

<212> DNA

<213> Hepatitis C virus

<400> 1

ggcatgcaag ctttaattaat t

21

<210> 2

<211> 68

<212> DNA

<213> Hepatitis C virus

<400> 2

ccggggaggc ctgcacgtga tcgagggcag acaccatcac caccatcact aatagttaat

60

taactgca

68

<210> 3

<211> 642

<212> DNA

<213> Hepatitis C virus

<220>

<221> CDS

<222> 1..639

<220>

<221> mat\_peptide

<222> 1..636

<400> 3

atg ccc ggt tgc tct ttc tct atc ttc ctc ttg gct tta ctg tcc tgt  
Met Pro Gly Cys Ser Phe Ser Ile Phe Leu Leu Ala Leu Leu Ser Cys  
1 5 10 15

48

ctg acc att cca gct tcc gct tat gag gtg cgc aac gtg tcc ggg atg  
Leu Thr Ile Pro Ala Ser Ala Tyr Glu Val Arg Asn Val Ser Gly Met  
20 25 30

96



Leu Thr Ile Pro Ala Ser Ala Tyr Glu Val Arg Asn Val Ser Gly Met  
                   20                  25                  30  
 Tyr His Val Thr Asn Asp Cys Ser Asn Ser Ser Ile Val Tyr Glu Ala  
                   35                  40                  45  
 Ala Asp Met Ile Met His Thr Pro Gly Cys Val Pro Cys Val Arg Glu  
                   50                  55                  60  
 Asn Asn Ser Ser Arg Cys Trp Val Ala Leu Thr Pro Thr Leu Ala Ala  
                   65                  70                  75                  80  
 Arg Asn Ala Ser Val Pro Thr Thr Thr Ile Arg Arg His Val Asp Leu  
                   85                  90                  95  
 Leu Val Gly Ala Ala Ala Leu Cys Ser Ala Met Tyr Val Gly Asp Leu  
                   100                  105                  110  
 Cys Gly Ser Val Phe Leu Val Ser Gln Leu Phe Thr Ile Ser Pro Arg  
                   115                  120                  125  
 Arg His Glu Thr Val Gln Asp Cys Asn Cys Ser Ile Tyr Pro Gly His  
                   130                  135                  140  
 Ile Thr Gly His Arg Met Ala Trp Asp Met Met Met Asn Trp Ser Pro  
                   145                  150                  155                  160  
 Thr Thr Ala Leu Val Val Ser Gln Leu Leu Arg Ile Pro Gln Ala Val  
                   165                  170                  175  
 Val Asp Met Val Ala Gly Ala His Trp Gly Val Leu Ala Gly Leu Ala  
                   180                  185                  190  
 Tyr Tyr Ser Met Val Gly Asn Trp Ala Lys Val Leu Ile Val Met Leu  
                   195                  200                  205  
 Leu Phe Ala Leu  
                   210

&lt;210&gt; 5

&lt;211&gt; 795

&lt;212&gt; DNA

&lt;213&gt; Hepatitis C virus

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; 1..792

&lt;220&gt;

&lt;221&gt; mat\_peptide

&lt;222&gt; 1..789

&lt;400&gt; 5

atg ttg ggt aag gtc atc gat acc ctt aca tgc ggc ttc gcc gac ctc  
 Met Leu Gly Lys Val Ile Asp Thr Leu Thr Cys Gly Phe Ala Asp Leu  
                   1                  5                  10                  15

48

gtg ggg tac att ccg ctc gtc ggc gcc ccc cta ggg ggc gct gcc agg  
 Val Gly Tyr Ile Pro Leu Val Gly Ala Pro Leu Gly Gly Ala Ala Arg

96

20	25	30	
gcc ctg gcg cat ggc gtc cgg gtt ctg gag gac ggc gtg aac tat gca Ala Leu Ala His Gly Val Arg Val Leu Glu Asp Gly Val Asn Tyr Ala			144
35	40	45	
aca ggg aat ttg ccc ggt tgc tct ttc tct atc ttc ctc ttg gct ttg Thr Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile Phe Leu Leu Ala Leu			192
50	55	60	
ctg tcc tgt ctg acc gtt cca gct tcc gct tat gaa gtg cgc aac gtg Leu Ser Cys Leu Thr Val Pro Ala Ser Ala Tyr Glu Val Arg Asn Val			240
65	70	75	80
tcc ggg atg tac cat gtc acg aac gac tgc tcc aac tca agc att gtg Ser Gly Met Tyr His Val Thr Asn Asp Cys Ser Asn Ser Ser Ile Val			288
85	90	95	
tat gag gca gcg gac atg atc atg cac acc ccc ggg tgc gtg ccc tgc Tyr Glu Ala Ala Asp Met Ile Met His Thr Pro Gly Cys Val Pro Cys			336
100	105	110	
gtt cgg gag aac aac tct tcc cgc tgc tgg gta gcg ctc acc ccc acg Val Arg Glu Asn Asn Ser Ser Arg Cys Trp Val Ala Leu Thr Pro Thr			384
115	120	125	
ctc gca gct agg aac gcc agc gtc ccc acc acg aca ata cga cgc cac Leu Ala Ala Arg Asn Ala Ser Val Pro Thr Thr Thr Ile Arg Arg His			432
130	135	140	
gtc gat ttg ctc gtt ggg gcg gct gct ttc tgt tcc gct atg tac gtg Val Asp Leu Leu Val Gly Ala Ala Phe Cys Ser Ala Met Tyr Val			480
145	150	155	160
ggg gac ctc tgc gga tct gtc ttc ctc gtc tcc cag ctg ttc acc atc Gly Asp Leu Cys Gly Ser Val Phe Leu Val Ser Gln Leu Phe Thr Ile			528
165	170	175	
tcg cct cgc cgg cat gag acg gtg cag gac tgc aat tgc tca atc tat Ser Pro Arg Arg His Glu Thr Val Gln Asp Cys Asn Cys Ser Ile Tyr			576
180	185	190	
ccc ggc cac ata acg ggt cac cgt atg gct tgg gat atg atg atg aac Pro Gly His Ile Thr Gly His Arg Met Ala Trp Asp Met Met Met Asn			624
195	200	205	
tgg tcg cct aca acg gcc ctg gtg gta tcg cag ctg ctc cgg atc cca Trp Ser Pro Thr Thr Ala Leu Val Val Ser Gln Leu Leu Arg Ile Pro			672
210	215	220	
caa gct gtc gtg gac atg gtg gcg ggg gcc cat tgg gga gtc ctg gcg Gln Ala Val Val Asp Met Val Ala Gly Ala His Trp Gly Val Leu Ala			720
225	230	235	240
ggt ctc gcc tac tat tcc atg gtg ggg aac tgg gct aag gtt ttg att Gly Leu Ala Tyr Tyr Ser Met Val Gly Asn Trp Ala Lys Val Leu Ile			768
245	250	255	
gtg atg cta ctc ttt gct ccc taatag Val Met Leu Leu Phe Ala Pro			795
260			

&lt;210&gt; 6



&lt;211&gt; 263

&lt;212&gt; PRT

&lt;213&gt; Hepatitis C virus

&lt;400&gt; 6

```

Met Leu Gly Lys Val Ile Asp Thr Leu Thr Cys Gly Phe Ala Asp Leu
 1             5             10             15
Val Gly Tyr Ile Pro Leu Val Gly Ala Pro Leu Gly Gly Ala Ala Arg
      20             25             30
Ala Leu Ala His Gly Val Arg Val Leu Glu Asp Gly Val Asn Tyr Ala
      35             40             45
Thr Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile Phe Leu Leu Ala Leu
      50             55             60
Leu Ser Cys Leu Thr Val Pro Ala Ser Ala Tyr Glu Val Arg Asn Val
      65             70             75             80
Ser Gly Met Tyr His Val Thr Asn Asp Cys Ser Asn Ser Ser Ile Val
      85             90             95
Tyr Glu Ala Ala Asp Met Ile Met His Thr Pro Gly Cys Val Pro Cys
      100             105             110
Val Arg Glu Asn Asn Ser Ser Arg Cys Trp Val Ala Leu Thr Pro Thr
      115             120             125
Leu Ala Ala Arg Asn Ala Ser Val Pro Thr Thr Thr Ile Arg Arg His
      130             135             140
Val Asp Leu Leu Val Gly Ala Ala Ala Phe Cys Ser Ala Met Tyr Val
      145             150             155             160
Gly Asp Leu Cys Gly Ser Val Phe Leu Val Ser Gln Leu Phe Thr Ile
      165             170             175
Ser Pro Arg Arg His Glu Thr Val Gln Asp Cys Asn Cys Ser Ile Tyr
      180             185             190
Pro Gly His Ile Thr Gly His Arg Met Ala Trp Asp Met Met Met Asn
      195             200             205
Trp Ser Pro Thr Thr Ala Leu Val Val Ser Gln Leu Leu Arg Ile Pro
      210             215             220
Gln Ala Val Val Asp Met Val Ala Gly Ala His Trp Gly Val Leu Ala
      225             230             235             240
Gly Leu Ala Tyr Tyr Ser Met Val Gly Asn Trp Ala Lys Val Leu Ile
      245             250             255
Val Met Leu Leu Phe Ala Pro
      260

```

&lt;210&gt; 7

<211> 633  
 <212> DNA  
 <213> Hepatitis C virus

<220>  
 <221> CDS  
 <222> 1..630

<220>  
 <221> mat\_peptide  
 <222> 1..627  
 <400> 7

atg ttg ggt aag gtc atc gat acc ctt acg tgc ggc ttc gcc gac ctc	48
Met Leu Gly Lys Val Ile Asp Thr Leu Thr Cys Gly Phe Ala Asp Leu	
1 5 10 15	
atg ggg tac att ccg ctc gtc ggc gcc ccc cta ggg ggt gct gcc aga	96
Met Gly Tyr Ile Pro Leu Val Gly Ala Pro Leu Gly Gly Ala Ala Arg	
20 25 30	
gcc ctg gcg cat ggc gtc cgg gtt ctg gaa gac ggc gtg aac tat gca	144
Ala Leu Ala His Gly Val Arg Val Leu Glu Asp Gly Val Asn Tyr Ala	
35 40 45	
aca ggg aat ttg cct ggt tgc tct ttc tct atc ttc ctc ttg gct tta	192
Thr Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile Phe Leu Leu Ala Leu	
50 55 60	
ctg tcc tgt ctg acc att cca gct tcc gct tat gag gtg cgc aac gtg	240
Leu Ser Cys Leu Thr Ile Pro Ala Ser Ala Tyr Glu Val Arg Asn Val	
65 70 75 80	
tcc ggg atg tac cat gtc acg aac gac tgc tcc aac tca agc att gtg	288
Ser Gly Met Tyr His Val Thr Asn Asp Cys Ser Asn Ser Ser Ile Val	
85 90 95	
tat gag gca gcg gac atg atc atg cac acc ccc ggg tgc gtg ccc tgc	336
Tyr Glu Ala Ala Asp Met Ile Met His Thr Pro Gly Cys Val Pro Cys	
100 105 110	
gtt cgg gag aac aac tct tcc cgc tgc tgg gta gcg ctc acc ccc acg	384
Val Arg Glu Asn Asn Ser Ser Arg Cys Trp Val Ala Leu Thr Pro Thr	
115 120 125	
ctc gca gct agg aac gcc agc gtc ccc act acg aca ata cga cgc cac	432
Leu Ala Ala Arg Asn Ala Ser Val Pro Thr Thr Thr Ile Arg Arg His	
130 135 140	
gtc gat ttg ctc gtt ggg gcg gct gct ttc tgt tcc gct atg tac gtg	480
Val Asp Leu Leu Val Gly Ala Ala Ala Phe Cys Ser Ala Met Tyr Val	
145 150 155 160	
ggg gat ctc tgc gga tct gtc ttc ctc gtc tcc cag ctg ttc acc atc	528
Gly Asp Leu Cys Gly Ser Val Phe Leu Val Ser Gln Leu Phe Thr Ile	
165 170 175	
tcg cct cgc cgg cat gag acg gtg cag gac tgc aat tgc tca atc tat	576
Ser Pro Arg Arg His Glu Thr Val Gln Asp Cys Asn Cys Ser Ile Tyr	
180 185 190	

tgg taatag	633
Trp	
210	

<213> Hepatitis C virus

Pro Gly His Ile Thr Gly His Arg Met Ala Trp Asp Met Met Met Asn  
195 200 205

Trp

&lt;210&gt; 9

&lt;211&gt; 483

&lt;212&gt; DNA

&lt;213&gt; Hepatitis C virus

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; 1..480

&lt;220&gt;

&lt;221&gt; mat\_peptide

&lt;222&gt; 1..477

&lt;400&gt; 9

```

atg ccc ggt tgc tct ttc tct atc ttc ctc ttg gcc ctg ctg tcc tgt      48
Met Pro Gly Cys Ser Phe Ser Ile Phe Leu Leu Ala Leu Leu Ser Cys
  1              5              10              15

ctg acc ata cca gct tcc gct tat gaa gtg cgc aac gtg tcc ggg gtg      96
Leu Thr Ile Pro Ala Ser Ala Tyr Glu Val Arg Asn Val Ser Gly Val
              20              25              30

tac cat gtc acg aac gac tgc tcc aac tca agc ata gtg tat gag gca     144
Tyr His Val Thr Asn Asp Cys Ser Asn Ser Ser Ile Val Tyr Glu Ala
              35              40              45

gcg gac atg atc atg cac acc ccc ggg tgc gtg ccc tgc gtt cgg gag     192
Ala Asp Met Ile Met His Thr Pro Gly Cys Val Pro Cys Val Arg Glu
              50              55              60

ggc aac tcc tcc cgt tgc tgg gtg gcg ctc act ccc acg ctc gcg gcc     240
Gly Asn Ser Ser Arg Cys Trp Val Ala Leu Thr Pro Thr Leu Ala Ala
              65              70              75              80

agg aac gcc agc gtc ccc aca acg aca ata cga cgc cac gtc gat ttg     288
Arg Asn Ala Ser Val Pro Thr Thr Thr Ile Arg Arg His Val Asp Leu
              85              90              95

ctc gtt ggg gct gct gct ttc tgt tcc gct atg tac gtg ggg gat ctc     336
Leu Val Gly Ala Ala Ala Phe Cys Ser Ala Met Tyr Val Gly Asp Leu
              100              105              110

tgc gga tct gtt ttc ctt gtt tcc cag ctg ttc acc ttc tca cct cgc     384
Cys Gly Ser Val Phe Leu Val Ser Gln Leu Phe Thr Phe Ser Pro Arg
              115              120              125

cgg cat caa aca gta cag gac tgc aac tgc tca atc tat ccc ggc cat     432
Arg His Gln Thr Val Gln Asp Cys Asn Cys Ser Ile Tyr Pro Gly His
              130              135              140

gta tca ggt cac cgc atg gct tgg gat atg atg atg aac tgg tcc taatag   483
Val Ser Gly His Arg Met Ala Trp Asp Met Met Met Asn Trp Ser
145              150              155              160

```

&lt;210&gt; 10

&lt;211&gt; 159

&lt;212&gt; PRT

&lt;213&gt; Hepatitis C virus

&lt;400&gt; 10

Met Pro Gly Cys Ser Phe Ser Ile Phe Leu Leu Ala Leu Leu Ser Cys  
 1 5 10 15  
 Leu Thr Ile Pro Ala Ser Ala Tyr Glu Val Arg Asn Val Ser Gly Val  
 20 25 30  
 Tyr His Val Thr Asn Asp Cys Ser Asn Ser Ser Ile Val Tyr Glu Ala  
 35 40 45  
 Ala Asp Met Ile Met His Thr Pro Gly Cys Val Pro Cys Val Arg Glu  
 50 55 60  
 Gly Asn Ser Ser Arg Cys Trp Val Ala Leu Thr Pro Thr Leu Ala Ala  
 65 70 75 80  
 Arg Asn Ala Ser Val Pro Thr Thr Thr Ile Arg Arg His Val Asp Leu  
 85 90 95  
 Leu Val Gly Ala Ala Ala Phe Cys Ser Ala Met Tyr Val Gly Asp Leu  
 100 105 110  
 Cys Gly Ser Val Phe Leu Val Ser Gln Leu Phe Thr Phe Ser Pro Arg  
 115 120 125  
 Arg His Gln Thr Val Gln Asp Cys Asn Cys Ser Ile Tyr Pro Gly His  
 130 135 140  
 Val Ser Gly His Arg Met Ala Trp Asp Met Met Met Asn Trp Ser  
 145 150 155

&lt;210&gt; 11

&lt;211&gt; 480

&lt;212&gt; DNA

&lt;213&gt; Hepatitis C virus

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; 1..477

&lt;220&gt;

&lt;221&gt; mat\_peptide

&lt;222&gt; 1..474

&lt;400&gt; 11

atg tcc ggt tgc tct ttc tct atc ttc ctc ttg gcc ctg ctg tcc tgt 48  
 Met Ser Gly Cys Ser Phe Ser Ile Phe Leu Leu Ala Leu Leu Ser Cys  
 1 5 10 15  
 ctg acc ata cca gct tcc gct tat gaa gtg cgc aac gtg tcc ggg gtg 96  
 Leu Thr Ile Pro Ala Ser Ala Tyr Glu Val Arg Asn Val Ser Gly Val  
 20 25 30  
 tac cat gtc acg aac gac tgc tcc aac tca agc ata gtg tat gag gca 144  
 Tyr His Val Thr Asn Asp Cys Ser Asn Ser Ser Ile Val Tyr Glu Ala  
 35 40 45  
 gcg gac atg atc atg cac acc ccc ggg tgc gtg ccc tgc gtt cgg gag 192  
 Ala Asp Met Ile Met His Thr Pro Gly Cys Val Pro Cys Val Arg Glu

50 55 60

ggc aac tcc tcc cgt tgc tgg gtg gcg ctc act ccc acg ctc gcg gcc 240  
 Gly Asn Ser Ser Arg Cys Trp Val Ala Leu Thr Pro Thr Leu Ala Ala  
 65 70 75 80

agg aac gcc agc gtc ccc aca acg aca ata cga cgc cac gtc gat ttg 288  
 Arg Asn Ala Ser Val Pro Thr Thr Thr Ile Arg Arg His Val Asp Leu  
 85 90 95

ctc gtt ggg gct gct gct ttc tgt tcc gct atg tac gtg ggg gat ctc 336  
 Leu Val Gly Ala Ala Phe Cys Ser Ala Met Tyr Val Gly Asp Leu  
 100 105 110

tgc gga tct gtt ttc ctt gtt tcc cag ctg ttc acc ttc tca cct cgc 384  
 Cys Gly Ser Val Phe Leu Val Ser Gln Leu Phe Thr Phe Ser Pro Arg  
 115 120 125

cgg cat caa aca gta cag gac tgc aac tgc tca atc tat ccc ggc cat 432  
 Arg His Gln Thr Val Gln Asp Cys Asn Cys Ser Ile Tyr Pro Gly His  
 130 135 140

gta tca ggt cac cgc atg gct tgg gat atg atg atg aac tgg taatag 480  
 Val Ser Gly His Arg Met Ala Trp Asp Met Met Met Asn Trp  
 145 150 155

<210> 12

<211> 158

<212> PRT

<213> Hepatitis C virus

<400> 12

Met Ser Gly Cys Ser Phe Ser Ile Phe Leu Leu Ala Leu Leu Ser Cys  
 1 5 10 15

Leu Thr Ile Pro Ala Ser Ala Tyr Glu Val Arg Asn Val Ser Gly Val  
 20 25 30

Tyr His Val Thr Asn Asp Cys Ser Asn Ser Ser Ile Val Tyr Glu Ala  
 35 40 45

Ala Asp Met Ile Met His Thr Pro Gly Cys Val Pro Cys Val Arg Glu  
 50 55 60

Gly Asn Ser Ser Arg Cys Trp Val Ala Leu Thr Pro Thr Leu Ala Ala  
 65 70 75 80

Arg Asn Ala Ser Val Pro Thr Thr Thr Ile Arg Arg His Val Asp Leu  
 85 90 95

Leu Val Gly Ala Ala Ala Phe Cys Ser Ala Met Tyr Val Gly Asp Leu  
 100 105 110

Cys Gly Ser Val Phe Leu Val Ser Gln Leu Phe Thr Phe Ser Pro Arg  
 115 120 125

Arg His Gln Thr Val Gln Asp Cys Asn Cys Ser Ile Tyr Pro Gly His  
 130 135 140

Val Ser Gly His Arg Met Ala Trp Asp Met Met Met Asn Trp  
 145 150 155

<210> 13

<211> 636

<212> DNA

<213> Hepatitis C virus

<220>

<221> CDS

<222> 1..633

<220>

<221> mat\_peptide

<222> 1..630

<400> 13

atg ctg ggt aag gcc atc gat acc ctt acg tgc ggc ttc gcc gac ctc	48
Met Leu Gly Lys Ala Ile Asp Thr Leu Thr Cys Gly Phe Ala Asp Leu	
1 5 10 15	
gtg ggg tac att ccg ctc gtc ggc gcc ccc cta ggg ggc gct gcc agg	96
Val Gly Tyr Ile Pro Leu Val Gly Ala Pro Leu Gly Gly Ala Ala Arg	
20 25 30	
gcc ctg gcg cat ggc gtc cgg gtt ctg gaa gac ggc gtg aac tat gca	144
Ala Leu Ala His Gly Val Arg Val Leu Glu Asp Gly Val Asn Tyr Ala	
35 40 45	
aca ggg aat ttg cct ggt tgc tct ttc tct atc ttc ctc ttg gct tta	192
Thr Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile Phe Leu Leu Ala Leu	
50 55 60	
ctg tcc tgt cta acc att cca gct tcc gct tac gag gtg cgc aac gtg	240
Leu Ser Cys Leu Thr Ile Pro Ala Ser Ala Tyr Glu Val Arg Asn Val	
65 70 75 80	
tcc ggg atg tac cat gtc acg aac gac tgc tcc aac tca agc att gtg	288
Ser Gly Met Tyr His Val Thr Asn Asp Cys Ser Asn Ser Ser Ile Val	
85 90 95	
tat gag gca gcg gac atg atc atg cac acc ccc ggg tgc gtg ccc tgc	336
Tyr Glu Ala Ala Asp Met Ile Met His Thr Pro Gly Cys Val Pro Cys	
100 105 110	
gtt cgg gag aac aac tct tcc cgc tgc tgg gta gcg ctc acc ccc acg	384
Val Arg Glu Asn Asn Ser Ser Arg Cys Trp Val Ala Leu Thr Pro Thr	
115 120 125	
ctc gcg gct agg aac gcc agc atc ccc act aca aca ata cga cgc cac	432
Leu Ala Ala Arg Asn Ala Ser Ile Pro Thr Thr Thr Ile Arg Arg His	
130 135 140	
gtc gat ttg ctc gtt ggg gcg gct gct ttc tgt tcc gct atg tac gtg	480
Val Asp Leu Leu Val Gly Ala Ala Ala Phe Cys Ser Ala Met Tyr Val	
145 150 155 160	
ggg gat ctc tgc gga tct gtc ttc ctc gtc tcc cag ctg ttc acc atc	528
Gly Asp Leu Cys Gly Ser Val Phe Leu Val Ser Gln Leu Phe Thr Ile	
165 170 175	

tcg cct cgc cgg cat gag acg gtg cag gac tgc aat tgc tca atc tat 576  
 Ser Pro Arg Arg His Glu Thr Val Gln Asp Cys Asn Cys Ser Ile Tyr  
                   180                  185                  190

ccc ggc cac ata acg ggt cac cgt atg gct tgg gat atg atg atg aac 624  
 Pro Gly His Ile Thr Gly His Arg Met Ala Trp Asp Met Met Met Asn  
                   195                  200                  205

tgg tac taatag 640  
 Trp Tyr  
           210

<210> 14

<211> 210

<212> PRT

<213> Hepatitis C virus

<400> 14

Met Leu Gly Lys Ala Ile Asp Thr Leu Thr Cys Gly Phe Ala Asp Leu  
   1                  5                  10                  15

Val Gly Tyr Ile Pro Leu Val Gly Ala Pro Leu Gly Gly Ala Ala Arg  
           20                  25                  30

Ala Leu Ala His Gly Val Arg Val Leu Glu Asp Gly Val Asn Tyr Ala  
           35                  40                  45

Thr Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile Phe Leu Leu Ala Leu  
           50                  55                  60

Leu Ser Cys Leu Thr Ile Pro Ala Ser Ala Tyr Glu Val Arg Asn Val  
   65                  70                  75                  80

Ser Gly Met Tyr His Val Thr Asn Asp Cys Ser Asn Ser Ser Ile Val  
                   85                  90                  95

Tyr Glu Ala Ala Asp Met Ile Met His Thr Pro Gly Cys Val Pro Cys  
           100                  105                  110

Val Arg Glu Asn Asn Ser Ser Arg Cys Trp Val Ala Leu Thr Pro Thr  
           115                  120                  125

Leu Ala Ala Arg Asn Ala Ser Ile Pro Thr Thr Thr Ile Arg Arg His  
           130                  135                  140

Val Asp Leu Leu Val Gly Ala Ala Ala Phe Cys Ser Ala Met Tyr Val  
   145                  150                  155                  160

Gly Asp Leu Cys Gly Ser Val Phe Leu Val Ser Gln Leu Phe Thr Ile  
           165                  170                  175

Ser Pro Arg Arg His Glu Thr Val Gln Asp Cys Asn Cys Ser Ile Tyr  
           180                  185                  190

Pro Gly His Ile Thr Gly His Arg Met Ala Trp Asp Met Met Met Asn  
           195                  200                  205

Trp Tyr



210  
<210> 15  
  
<211> 26  
<212> DNA  
<213> Hepatitis C virus  
<400> 15  
  
atgcccggtt gctctttctc tatctt 26  
  
<210> 16  
  
<211> 26  
<212> DNA  
<213> Hepatitis C virus  
  
<400> 16  
  
atgttgggta aggtcatcga taccct 26  
  
<210> 17  
  
<211> 30  
<212> DNA  
<213> Hepatitis C virus  
  
<220>  
<221> misc\_feature  
<223>/note="antisense"  
  
<400> 17  
  
ctattaggac cagttcatca tcatatccca 30  
  
<210> 18  
  
<211> 27  
<212> DNA  
<213> Hepatitis C virus  
  
<400> 18  
  
ctattaccag ttcacatca tatccca 27  
  
<210> 19  
  
<211> 36  
<212> DNA  
<213> Hepatitis C virus  
<400> 19  
  
atacgacgcc acgtcgattc ccagctgttc accatc 36  
  
<210> 20

<211> 36  
 <212> DNA  
 <213> Hepatitis C virus  
 <220>  
 <221> misc\_feature  
 <223>/note="antisense"

<400> 20

gatggtgaac agctgggaat cgacgtggcg tcgtat

36

<210> 21

<211> 723  
 <212> DNA  
 <213> Hepatitis C virus

<220>  
 <221> CDS  
 <222> 1..720

<220>  
 <221> mat\_peptide  
 <222> 1..717  
 <400> 21

atg ttg ggt aag gtc atc gat acc ctt aca tgc ggc ttc gcc gac ctc  
 Met Leu Gly Lys Val Ile Asp Thr Leu Thr Cys Gly Phe Ala Asp Leu  
 1 5 10 15

48

gtg ggg tac att ccg ctc gtc ggc gcc ccc cta ggg ggc gct gcc agg  
 Val Gly Tyr Ile Pro Leu Val Gly Ala Pro Leu Gly Gly Ala Ala Arg  
 20 25 30

96

gcc ctg gcg cat ggc gtc cgg gtt ctg gag gac ggc gtg aac tat gca  
 Ala Leu Ala His Gly Val Arg Val Leu Glu Asp Gly Val Asn Tyr Ala  
 35 40 45

144

aca ggg aat ttg ccc ggt tgc tct ttc tct atc ttc ctc ttg gct ttg  
 Thr Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile Phe Leu Leu Ala Leu  
 50 55 60

192

ctg tcc tgt ctg acc gtt cca gct tcc gct tat gaa gtg cgc aac gtg  
 Leu Ser Cys Leu Thr Val Pro Ala Ser Ala Tyr Glu Val Arg Asn Val  
 65 70 75 80

240

tcc ggg atg tac cat gtc acg aac gac tgc tcc aac tca agc att gtg  
 Ser Gly Met Tyr His Val Thr Asn Asp Cys Ser Asn Ser Ser Ile Val  
 85 90 95

288

tat gag gca gcg gac atg atc atg cac acc ccc ggg tgc gtg ccc tgc  
 Tyr Glu Ala Ala Asp Met Ile Met His Thr Pro Gly Cys Val Pro Cys  
 100 105 110

336

gtt cgg gag aac aac tct tcc cgc tgc tgg gta gcg ctc acc ccc acg  
 Val Arg Glu Asn Asn Ser Ser Arg Cys Trp Val Ala Leu Thr Pro Thr  
 115 120 125

384

ctc gca gct agg aac gcc agc gtc ccc acc acg aca ata cga cgc cac  
 Leu Ala Ala Arg Asn Ala Ser Val Pro Thr Thr Thr Ile Arg Arg His

432

130 135 140  
 gtc gat tcc cag ctg ttc acc atc tcg cct cgc cgg cat gag acg gtg 480  
 Val Asp Ser Gln Leu Phe Thr Ile Ser Pro Arg Arg His Glu Thr Val  
 145 150 155 160  
 cag gac tgc aat tgc tca atc tat ccc ggc cac ata acg ggt cac cgt 528  
 Gln Asp Cys Asn Cys Ser Ile Tyr Pro Gly His Ile Thr Gly His Arg  
 165 170 175  
 atg gct tgg gat atg atg atg aac tgg tcg cct aca acg gcc ctg gtg 576  
 Met Ala Trp Asp Met Met Met Asn Trp Ser Pro Thr Thr Ala Leu Val  
 180 185 190  
 gta tcg cag ctg ctc cgg atc cca caa gct gtc gtg gac atg gtg gcg 624  
 Val Ser Gln Leu Leu Arg Ile Pro Gln Ala Val Val Asp Met Val Ala  
 195 200 205  
 ggg gcc cat tgg gga gtc ctg gcg ggt ctc gcc tac tat tcc atg gtg 672  
 Gly Ala His Trp Gly Val Leu Ala Gly Leu Ala Tyr Tyr Ser Met Val  
 210 215 220  
 ggg aac tgg gct aag gtt ttg att gtg atg cta ctc ttt gct ccc taatag 723  
 Gly Asn Trp Ala Lys Val Leu Ile Val Met Leu Leu Phe Ala Pro  
 225 230 235 240  
 <210> 22  
 <211> 239  
 <212> PRT  
 <213> Hepatitis C virus  
 <400> 22  
 Met Leu Gly Lys Val Ile Asp Thr Leu Thr Cys Gly Phe Ala Asp Leu  
 1 5 10 15  
 Val Gly Tyr Ile Pro Leu Val Gly Ala Pro Leu Gly Gly Ala Ala Arg  
 20 25 30  
 Ala Leu Ala His Gly Val Arg Val Leu Glu Asp Gly Val Asn Tyr Ala  
 35 40 45  
 Thr Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile Phe Leu Leu Ala Leu  
 50 55 60  
 Leu Ser Cys Leu Thr Val Pro Ala Ser Ala Tyr Glu Val Arg Asn Val  
 65 70 75 80  
 Ser Gly Met Tyr His Val Thr Asn Asp Cys Ser Asn Ser Ser Ile Val  
 85 90 95  
 Tyr Glu Ala Ala Asp Met Ile Met His Thr Pro Gly Cys Val Pro Cys  
 100 105 110  
 Val Arg Glu Asn Asn Ser Ser Arg Cys Trp Val Ala Leu Thr Pro Thr  
 115 120 125  
 Leu Ala Ala Arg Asn Ala Ser Val Pro Thr Thr Thr Ile Arg Arg His  
 130 135 140  
 Val Asp Ser Gln Leu Phe Thr Ile Ser Pro Arg Arg His Glu Thr Val

145                      150                      155                      160  
 Gln Asp Cys Asn Cys Ser Ile Tyr Pro Gly His Ile Thr Gly His Arg  
                                  165                      170                      175  
 Met Ala Trp Asp Met Met Met Asn Trp Ser Pro Thr Thr Ala Leu Val  
                                  180                      185                      190  
 Val Ser Gln Leu Leu Arg Ile Pro Gln Ala Val Val Asp Met Val Ala  
                                  195                      200                      205  
 Gly Ala His Trp Gly Val Leu Ala Gly Leu Ala Tyr Tyr Ser Met Val  
                                  210                      215                      220  
 Gly Asn Trp Ala Lys Val Leu Ile Val Met Leu Leu Phe Ala Pro  
                                  225                      230                      235

<210> 23

<211> 561

<212> DNA

<213> Hepatitis C virus

<220>

<221> CDS

<222> 1..558

<220>

<221> mat\_peptide

<222> 1..555

<400> 23

atg ttg ggt aag gtc atc gat acc ctt aca tgc ggc ttc gcc gac ctc                      48  
 Met Leu Gly Lys Val Ile Asp Thr Leu Thr Cys Gly Phe Ala Asp Leu  
                                  1                      5                      10                      15  
 gtg ggg tac att ccg ctc gtc ggc gcc ccc cta ggg ggc gct gcc agg                      96  
 Val Gly Tyr Ile Pro Leu Val Gly Ala Pro Leu Gly Gly Ala Ala Arg  
                                  20                      25                      30  
 gcc ctg gcg cat ggc gtc cgg gtt ctg gag gac ggc gtg aac tat gca                      144  
 Ala Leu Ala His Gly Val Arg Val Leu Glu Asp Gly Val Asn Tyr Ala  
                                  35                      40                      45  
 aca ggg aat ttg ccc ggt tgc tct ttc tct atc ttc ctc ttg gct ttg                      192  
 Thr Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile Phe Leu Leu Ala Leu  
                                  50                      55                      60  
 ctg tcc tgt ctg acc gtt cca gct tcc gct tat gaa gtg cgc aac gtg                      240  
 Leu Ser Cys Leu Thr Val Pro Ala Ser Ala Tyr Glu Val Arg Asn Val  
                                  65                      70                      75                      80  
 tcc ggg atg tac cat gtc acg aac gac tgc tcc aac tca agc att gtg                      288  
 Ser Gly Met Tyr His Val Thr Asn Asp Cys Ser Asn Ser Ser Ile Val  
                                  85                      90                      95  
 tat gag gca gcg gac atg atc atg cac acc ccc ggg tgc gtg ccc tgc                      336  
 Tyr Glu Ala Ala Asp Met Ile Met His Thr Pro Gly Cys Val Pro Cys  
                                  100                      105                      110  
 gtt cgg gag aac aac tct tcc cgc tgc tgg gta gcg ctc acc ccc acg                      384  
 Val Arg Glu Asn Asn Ser Ser Arg Cys Trp Val Ala Leu Thr Pro Thr

115	120	125	
ctc gca gct agg aac gcc agc gtc ccc acc acg aca ata cga cgc cac			432
Leu Ala Ala Arg Asn Ala Ser Val Pro Thr Thr Thr Ile Arg Arg His			
130	135	140	
gtc gat tcc cag ctg ttc acc atc tcg cct cgc cgg cat gag acg gtg			480
Val Asp Ser Gln Leu Phe Thr Ile Ser Pro Arg Arg His Glu Thr Val			
145	150	155	160
cag gac tgc aat tgc tca atc tat ccc ggc cac ata acg ggt cac cgt			528
Gln Asp Cys Asn Cys Ser Ile Tyr Pro Gly His Ile Thr Gly His Arg			
165	170	175	
atg gct tgg gat atg atg atg aac tgg taatag			561
Met Ala Trp Asp Met Met Met Asn Trp			
180	185		

&lt;210&gt; 24

&lt;211&gt; 185

&lt;212&gt; PRT

&lt;213&gt; Hepatitis C virus

&lt;400&gt; 24

Met Leu Gly Lys Val Ile Asp Thr Leu Thr Cys Gly Phe Ala Asp Leu			
1	5	10	15
Val Gly Tyr Ile Pro Leu Val Gly Ala Pro Leu Gly Gly Ala Ala Arg			
20	25	30	
Ala Leu Ala His Gly Val Arg Val Leu Glu Asp Gly Val Asn Tyr Ala			
35	40	45	
Thr Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile Phe Leu Leu Ala Leu			
50	55	60	
Leu Ser Cys Leu Thr Val Pro Ala Ser Ala Tyr Glu Val Arg Asn Val			
65	70	75	80
Ser Gly Met Tyr His Val Thr Asn Asp Cys Ser Asn Ser Ser Ile Val			
85	90	95	
Tyr Glu Ala Ala Asp Met Ile Met His Thr Pro Gly Cys Val Pro Cys			
100	105	110	
Val Arg Glu Asn Asn Ser Ser Arg Cys Trp Val Ala Leu Thr Pro Thr			
115	120	125	
Leu Ala Ala Arg Asn Ala Ser Val Pro Thr Thr Thr Ile Arg Arg His			
130	135	140	
Val Asp Ser Gln Leu Phe Thr Ile Ser Pro Arg Arg His Glu Thr Val			
145	150	155	160
Gln Asp Cys Asn Cys Ser Ile Tyr Pro Gly His Ile Thr Gly His Arg			
165	170	175	
Met Ala Trp Asp Met Met Met Asn Trp			
180	185		

&lt;210&gt; 25

&lt;211&gt; 606

&lt;212&gt; DNA

&lt;213&gt; Hepatitis C virus

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; 1..603

&lt;220&gt;

&lt;221&gt; mat\_peptide

&lt;222&gt; 1..600

&lt;400&gt; 25

atg ttg ggt aag gtc atc gat acc ctt aca tgc ggc ttc gcc gac ctc	48
Met Leu Gly Lys Val Ile Asp Thr Leu Thr Cys Gly Phe Ala Asp Leu	
1 5 10 15	
gtg ggg tac att ccg ctc gtc ggc gcc ccc cta ggg ggc gct gcc agg	96
Val Gly Tyr Ile Pro Leu Val Gly Ala Pro Leu Gly Gly Ala Ala Arg	
20 25 30	
gcc ctg gcg cat ggc gtc cgg gtt ctg gag gac ggc gtg aac tat gca	144
Ala Leu Ala His Gly Val Arg Val Leu Glu Asp Gly Val Asn Tyr Ala	
35 40 45	
aca ggg aat ttg ccc ggt tgc tct ttc tct atc ttc ctc ttg gct ttg	192
Thr Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile Phe Leu Leu Ala Leu	
50 55 60	
ctg tcc tgt ctg acc gtt cca gct tcc gct tat gaa gtg cgc aac gtg	240
Leu Ser Cys Leu Thr Val Pro Ala Ser Ala Tyr Glu Val Arg Asn Val	
65 70 75 80	
tcc ggg atg tac cat gtc acg aac gac tgc tcc aac tca agc att gtg	288
Ser Gly Met Tyr His Val Thr Asn Asp Cys Ser Asn Ser Ser Ile Val	
85 90 95	
tat gag gca gcg gac atg atc atg cac acc ccc ggg tgc gtg ccc tgc	336
Tyr Glu Ala Ala Asp Met Ile Met His Thr Pro Gly Cys Val Pro Cys	
100 105 110	
gtt cgg gag aac aac tct tcc cgc tgc tgg gta gcg ctc acc ccc acg	384
Val Arg Glu Asn Asn Ser Ser Arg Cys Trp Val Ala Leu Thr Pro Thr	
115 120 125	
ctc gca gct agg aac gcc agc gtc ccc acc acg aca ata cga cgc cac	432
Leu Ala Ala Arg Asn Ala Ser Val Pro Thr Thr Ile Arg Arg His	
130 135 140	
gtc gat tcc cag ctg ttc acc atc tcg cct cgc cgg cat gag acg gtg	480
Val Asp Ser Gln Leu Phe Thr Ile Ser Pro Arg Arg His Glu Thr Val	
145 150 155 160	
cag gac tgc aat tgc tca atc tat ccc ggc cac ata acg ggt cac cgt	528
Gln Asp Cys Asn Cys Ser Ile Tyr Pro Gly His Ile Thr Gly His Arg	
165 170 175	
atg gct tgg gat atg atg atg aac tgg tcg cct aca acg gcc ctg gtg	576
Met Ala Trp Asp Met Met Met Asn Trp Ser Pro Thr Thr Ala Leu Val	

180 185 190  
 gta tgc cag ctg ctc cgg atc ctc taatag 606  
 Val Ser Gln Leu Leu Arg Ile Leu  
 195 200  
 <210> 26  
 <211> 200  
 <212> PRT  
 <213> Hepatitis C virus  
 <400> 26  
 Met Leu Gly Lys Val Ile Asp Thr Leu Thr Cys Gly Phe Ala Asp Leu  
 1 5 10 15  
 Val Gly Tyr Ile Pro Leu Val Gly Ala Pro Leu Gly Gly Ala Ala Arg  
 20 25 30  
 Ala Leu Ala His Gly Val Arg Val Leu Glu Asp Gly Val Asn Tyr Ala  
 35 40 45  
 Thr Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile Phe Leu Leu Ala Leu  
 50 55 60  
 Leu Ser Cys Leu Thr Val Pro Ala Ser Ala Tyr Glu Val Arg Asn Val  
 65 70 75 80  
 Ser Gly Met Tyr His Val Thr Asn Asp Cys Ser Asn Ser Ser Ile Val  
 85 90 95  
 Tyr Glu Ala Ala Asp Met Ile Met His Thr Pro Gly Cys Val Pro Cys  
 100 105 110  
 Val Arg Glu Asn Asn Ser Ser Arg Cys Trp Val Ala Leu Thr Pro Thr  
 115 120 125  
 Leu Ala Ala Arg Asn Ala Ser Val Pro Thr Thr Thr Ile Arg Arg His  
 130 135 140  
 Val Asp Ser Gln Leu Phe Thr Ile Ser Pro Arg Arg His Glu Thr Val  
 145 150 155 160  
 Gln Asp Cys Asn Cys Ser Ile Tyr Pro Gly His Ile Thr Gly His Arg  
 165 170 175  
 Met Ala Trp Asp Met Met Met Asn Trp Ser Pro Thr Thr Ala Leu Val  
 180 185 190  
 Val Ser Gln Leu Leu Arg Ile Leu  
 195 200  
 <210> 27  
 <211> 636  
 <212> DNA  
 <213> Hepatitis C virus  
 <220>  
 <221> CDS  
 <222> 1..633

&lt;220&gt;

&lt;221&gt; mat\_peptide

&lt;222&gt; 1..630

&lt;400&gt; 27

atg ttg ggt aag gtc atc gat acc ctt aca tgc ggc ttc gcc gac ctc	48
Met Leu Gly Lys Val Ile Asp Thr Leu Thr Cys Gly Phe Ala Asp Leu	
1 5 10 15	
gtg ggg tac att ccg ctc gtc ggc gcc ccc cta ggg ggc gct gcc agg	96
Val Gly Tyr Ile Pro Leu Val Gly Ala Pro Leu Gly Gly Ala Ala Arg	
20 25 30	
gcc ctg gcg cat ggc gtc cgg gtt ctg gag gac ggc gtg aac tat gca	144
Ala Leu Ala His Gly Val Arg Val Leu Glu Asp Gly Val Asn Tyr Ala	
35 40 45	
aca ggg aat ttg ccc ggt tgc tct ttc tct atc ttc ctc ttg gct ttg	192
Thr Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile Phe Leu Leu Ala Leu	
50 55 60	
ctg tcc tgt ctg acc gtt cca gct tcc gct tat gaa gtg cgc aac gtg	240
Leu Ser Cys Leu Thr Val Pro Ala Ser Ala Tyr Glu Val Arg Asn Val	
65 70 75 80	
tcc ggg atg tac cat gtc acg aac gac tgc tcc aac tca agc att gtg	288
Ser Gly Met Tyr His Val Thr Asn Asp Cys Ser Asn Ser Ser Ile Val	
85 90 95	
tat gag gca gcg gac atg atc atg cac acc ccc ggg tgc gtg ccc tgc	336
Tyr Glu Ala Ala Asp Met Ile Met His Thr Pro Gly Cys Val Pro Cys	
100 105 110	
gtt cgg gag aac aac tct tcc cgc tgc tgg gta gcg ctc acc ccc acg	384
Val Arg Glu Asn Asn Ser Ser Arg Cys Trp Val Ala Leu Thr Pro Thr	
115 120 125	
ctc gca gct agg aac gcc agc gtc ccc acc acg aca ata cga cgc cac	432
Leu Ala Ala Arg Asn Ala Ser Val Pro Thr Thr Thr Ile Arg Arg His	
130 135 140	
gtc gat tcc cag ctg ttc acc atc tcg cct cgc cgg cat gag acg gtg	480
Val Asp Ser Gln Leu Phe Thr Ile Ser Pro Arg Arg His Glu Thr Val	
145 150 155 160	
cag gac tgc aat tgc tca atc tat ccc ggc cac ata acg ggt cac cgt	528
Gln Asp Cys Asn Cys Ser Ile Tyr Pro Gly His Ile Thr Gly His Arg	
165 170 175	
atg gct tgg gat atg atg atg aac tgg tcg cct aca acg gcc ctg gtg	576
Met Ala Trp Asp Met Met Met Asn Trp Ser Pro Thr Thr Ala Leu Val	
180 185 190	
gta tcg cag ctg ctc cgg atc gtg atc gag ggc aga cac cat cac cac	624
Val Ser Gln Leu Leu Arg Ile Val Ile Glu Gly Arg His His His His	
195 200 205	
cat cac taatag	636
His His	
210	



&lt;210&gt; 28

&lt;211&gt; 210

&lt;212&gt; PRT

&lt;213&gt; Hepatitis C virus

&lt;400&gt; 28

Met Leu Gly Lys Val Ile Asp Thr Leu Thr Cys Gly Phe Ala Asp Leu  
 1 5 10 15  
 Val Gly Tyr Ile Pro Leu Val Gly Ala Pro Leu Gly Gly Ala Ala Arg  
 20 25 30  
 Ala Leu Ala His Gly Val Arg Val Leu Glu Asp Gly Val Asn Tyr Ala  
 35 40 45  
 Thr Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile Phe Leu Leu Ala Leu  
 50 55 60  
 Leu Ser Cys Leu Thr Val Pro Ala Ser Ala Tyr Glu Val Arg Asn Val  
 65 70 75 80  
 Ser Gly Met Tyr His Val Thr Asn Asp Cys Ser Asn Ser Ser Ile Val  
 85 90 95  
 Tyr Glu Ala Ala Asp Met Ile Met His Thr Pro Gly Cys Val Pro Cys  
 100 105 110  
 Val Arg Glu Asn Asn Ser Ser Arg Cys Trp Val Ala Leu Thr Pro Thr  
 115 120 125  
 Leu Ala Ala Arg Asn Ala Ser Val Pro Thr Thr Thr Ile Arg Arg His  
 130 135 140  
 Val Asp Ser Gln Leu Phe Thr Ile Ser Pro Arg Arg His Glu Thr Val  
 145 150 155 160  
 Gln Asp Cys Asn Cys Ser Ile Tyr Pro Gly His Ile Thr Gly His Arg  
 165 170 175  
 Met Ala Trp Asp Met Met Met Asn Trp Ser Pro Thr Thr Ala Leu Val  
 180 185 190  
 Val Ser Gln Leu Leu Arg Ile Val Ile Glu Gly Arg His His His His  
 195 200 205

His His  
 210

&lt;210&gt; 29

&lt;211&gt; 630

&lt;212&gt; DNA

&lt;213&gt; Hepatitis C virus

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; 1..627

&lt;220&gt;

&lt;221&gt; mat\_peptide

&lt;222&gt; 1..624

&lt;400&gt; 29

atg ggt aag gtc atc gat acc ctt acg tgc gga ttc gcc gat ctc atg	48
Met Gly Lys Val Ile Asp Thr Leu Thr Cys Gly Phe Ala Asp Leu Met	
1 5 10 15	
ggg tac atc ccg ctc gtc ggc gct ccc gta gga ggc gtc gca aga gcc	96
Gly Tyr Ile Pro Leu Val Gly Ala Pro Val Gly Gly Val Ala Arg Ala	
20 25 30	
ctt gcg cat ggc gtg agg gcc ctt gaa gac ggg ata aat ttc gca aca	144
Leu Ala His Gly Val Arg Ala Leu Glu Asp Gly Ile Asn Phe Ala Thr	
35 40 45	
ggg aat ttg ccc ggt tgc tcc ttt tct att ttc ctt ctc gct ctg ttc	192
Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile Phe Leu Leu Ala Leu Phe	
50 55 60	
tct tgc tta att cat cca gca gct agt cta gag tgg cgg aat acg tct	240
Ser Cys Leu Ile His Pro Ala Ala Ser Leu Glu Trp Arg Asn Thr Ser	
65 70 75 80	
ggc ctc tat gtc ctt acc aac gac tgt tcc aat agc agt att gtg tac	288
Gly Leu Tyr Val Leu Thr Asn Asp Cys Ser Asn Ser Ser Ile Val Tyr	
85 90 95	
gag gcc gat gac gtt att ctg cac aca ccc ggc tgc ata cct tgt gtc	336
Glu Ala Asp Asp Val Ile Leu His Thr Pro Gly Cys Ile Pro Cys Val	
100 105 110	
cag gac ggc aat aca tcc acg tgc tgg acc cca gtg aca cct aca gtg	384
Gln Asp Gly Asn Thr Ser Thr Cys Trp Thr Pro Val Thr Pro Thr Val	
115 120 125	
gca gtc aag tac gtc gga gca acc acc gct tcg ata cgc agt cat gtg	432
Ala Val Lys Tyr Val Gly Ala Thr Thr Ala Ser Ile Arg Ser His Val	
130 135 140	
gac cta tta gtg ggc gcg gcc acg atg tgc tct gcg ctc tac gtg ggt	480
Asp Leu Leu Val Gly Ala Ala Thr Met Cys Ser Ala Leu Tyr Val Gly	
145 150 155 160	
gac atg tgt ggg gct gtc ttc ctc gtg gga caa gcc ttc acg ttc aga	528
Asp Met Cys Gly Ala Val Phe Leu Val Gly Gln Ala Phe Thr Phe Arg	
165 170 175	
cct cgt cgc cat caa acg gtc cag acc tgt aac tgc tcg ctg tac cca	576
Pro Arg Arg His Gln Thr Val Gln Thr Cys Asn Cys Ser Leu Tyr Pro	
180 185 190	
ggc cat ctt tca gga cat cga atg gct tgg gat atg atg atg aac tgg	624
Gly His Leu Ser Gly His Arg Met Ala Trp Asp Met Met Met Asn Trp	
195 200 205	
taatag	634
<210> 30	
<211> 208	
<212> PRT	
<213> Hepatitis C virus	

&lt;400&gt; 30

Met Gly Lys Val Ile Asp Thr Leu Thr Cys Gly Phe Ala Asp Leu Met  
 1 5 10 15

Gly Tyr Ile Pro Leu Val Gly Ala Pro Val Gly Gly Val Ala Arg Ala  
 20 25 30

Leu Ala His Gly Val Arg Ala Leu Glu Asp Gly Ile Asn Phe Ala Thr  
 35 40 45

Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile Phe Leu Leu Ala Leu Phe  
 50 55 60

Ser Cys Leu Ile His Pro Ala Ala Ser Leu Glu Trp Arg Asn Thr Ser  
 65 70 75 80

Gly Leu Tyr Val Leu Thr Asn Asp Cys Ser Asn Ser Ser Ile Val Tyr  
 85 90 95

Glu Ala Asp Asp Val Ile Leu His Thr Pro Gly Cys Ile Pro Cys Val  
 100 105 110

Gln Asp Gly Asn Thr Ser Thr Cys Trp Thr Pro Val Thr Pro Thr Val  
 115 120 125

Ala Val Lys Tyr Val Gly Ala Thr Thr Ala Ser Ile Arg Ser His Val  
 130 135 140

Asp Leu Leu Val Gly Ala Ala Thr Met Cys Ser Ala Leu Tyr Val Gly  
 145 150 155 160

Asp Met Cys Gly Ala Val Phe Leu Val Gly Gln Ala Phe Thr Phe Arg  
 165 170 175

Pro Arg Arg His Gln Thr Val Gln Thr Cys Asn Cys Ser Leu Tyr Pro  
 180 185 190

Gly His Leu Ser Gly His Arg Met Ala Trp Asp Met Met Met Asn Trp  
 195 200 205

&lt;210&gt; 31

&lt;211&gt; 630

&lt;212&gt; DNA

&lt;213&gt; Hepatitis C virus

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; 1..627

&lt;220&gt;

&lt;221&gt; mat\_peptide

&lt;222&gt; 1..624

&lt;400&gt; 31

atg ggt aag gtc atc gat acc cta acg tgc gga ttc gcc gat ctc atg  
 Met Gly Lys Val Ile Asp Thr Leu Thr Cys Gly Phe Ala Asp Leu Met  
 1 5 10 15

ggg tat atc ccg ctc gta ggc ggc ccc att ggg ggc gtc gca agg gct 96  
 Gly Tyr Ile Pro Leu Val Gly Gly Pro Ile Gly Gly Val Ala Arg Ala  
                   20                                  25                                  30

ctc gca cac ggt gtg agg gtc ctt gag gac ggg gta aac tat gca aca 144  
 Leu Ala His Gly Val Arg Val Leu Glu Asp Gly Val Asn Tyr Ala Thr  
                   35                                  40                                  45

ggg aat tta ccc ggt tgc tct ttc tct atc ttt att ctt gct ctt ctc 192  
 Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile Phe Ile Leu Ala Leu Leu  
                   50                                  55                                  60

tcg tgt ctg acc gtt ccg gcc tct gca gtt ccc tac cga aat gcc tct 240  
 Ser Cys Leu Thr Val Pro Ala Ser Ala Val Pro Tyr Arg Asn Ala Ser  
                   65                                  70                                  75                                  80

ggg att tat cat gtt acc aat gat tgc cca aac tct tcc ata gtc tat 288  
 Gly Ile Tyr His Val Thr Asn Asp Cys Pro Asn Ser Ser Ile Val Tyr  
                   85                                  90                                  95

gag gca gat aac ctg atc cta cac gca cct ggt tgc gtg cct tgt gtc 336  
 Glu Ala Asp Asn Leu Ile Leu His Ala Pro Gly Cys Val Pro Cys Val  
                   100                                  105                                  110

atg aca ggt aat gtg agt aga tgc tgg gtc caa att acc cct aca ctg 384  
 Met Thr Gly Asn Val Ser Arg Cys Trp Val Gln Ile Thr Pro Thr Leu  
                   115                                  120                                  125

tca gcc ccg agc ctc gga gca gtc acg gct cct ctt cgg aga gcc gtt 432  
 Ser Ala Pro Ser Leu Gly Ala Val Thr Ala Pro Leu Arg Arg Ala Val  
                   130                                  135                                  140

gac tac cta gcg gga ggg gct gcc ctc tgc tcc gcg tta tac gta gga 480  
 Asp Tyr Leu Ala Gly Gly Ala Ala Leu Cys Ser Ala Leu Tyr Val Gly  
                   145                                  150                                  155                                  160

gac gcg tgt ggg gca cta ttc ttg gta ggc caa atg ttc acc tat agg 528  
 Asp Ala Cys Gly Ala Leu Phe Leu Val Gly Gln Met Phe Thr Tyr Arg  
                   165                                  170                                  175

cct cgc cag cac gct acg gtg cag aac tgc aac tgt tcc att tac agt 576  
 Pro Arg Gln His Ala Thr Val Gln Asn Cys Asn Cys Ser Ile Tyr Ser  
                   180                                  185                                  190

ggc cat gtt acc ggc cac cgg atg gca tgg gat atg atg atg aac tgg 624  
 Gly His Val Thr Gly His Arg Met Ala Trp Asp Met Met Met Asn Trp  
                   195                                  200                                  205

taatag 630

&lt;210&gt; 32

&lt;211&gt; 208

&lt;212&gt; PRT

&lt;213&gt; Hepatitis C virus

&lt;400&gt; 32

Met Gly Lys Val Ile Asp Thr Leu Thr Cys Gly Phe Ala Asp Leu Met  
           1                                  5                                  10                                  15

Gly Tyr Ile Pro Leu Val Gly Gly Pro Ile Gly Gly Val Ala Arg Ala  
                   20                  25                  30  
 Leu Ala His Gly Val Arg Val Leu Glu Asp Gly Val Asn Tyr Ala Thr  
                   35                  40                  45  
 Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile Phe Ile Leu Ala Leu Leu  
                   50                  55                  60  
 Ser Cys Leu Thr Val Pro Ala Ser Ala Val Pro Tyr Arg Asn Ala Ser  
                   65                  70                  75                  80  
 Gly Ile Tyr His Val Thr Asn Asp Cys Pro Asn Ser Ser Ile Val Tyr  
                   85                  90                  95  
 Glu Ala Asp Asn Leu Ile Leu His Ala Pro Gly Cys Val Pro Cys Val  
                   100                  105                  110  
 Met Thr Gly Asn Val Ser Arg Cys Trp Val Gln Ile Thr Pro Thr Leu  
                   115                  120                  125  
 Ser Ala Pro Ser Leu Gly Ala Val Thr Ala Pro Leu Arg Arg Ala Val  
                   130                  135                  140  
 Asp Tyr Leu Ala Gly Gly Ala Ala Leu Cys Ser Ala Leu Tyr Val Gly  
                   145                  150                  155                  160  
 Asp Ala Cys Gly Ala Leu Phe Leu Val Gly Gln Met Phe Thr Tyr Arg  
                   165                  170                  175  
 Pro Arg Gln His Ala Thr Val Gln Asn Cys Asn Cys Ser Ile Tyr Ser  
                   180                  185                  190  
 Gly His Val Thr Gly His Arg Met Ala Trp Asp Met Met Met Asn Trp  
                   195                  200                  205

&lt;210&gt; 33

&lt;211&gt; 23

&lt;212&gt; DNA

&lt;213&gt; Hepatitis C virus

&lt;400&gt; 33

tgggatatga tgatgaactg gtc

23

&lt;210&gt; 34

&lt;211&gt; 30

&lt;212&gt; DNA

&lt;213&gt; Hepatitis C virus

&lt;400&gt; 34

ctattatggt ggtaagccac agagcaggag

30

&lt;210&gt; 35

&lt;211&gt; 1476

&lt;212&gt; DNA

&lt;213&gt; Hepatitis C virus

<220>  
 <221> CDS  
 <222> 1..1473

<220>  
 <221> mat\_peptide  
 <222> 1..1470  
 <400> 35

tgg gat atg atg atg aac tgg tcg cct aca acg gcc ctg gtg gta tcg	48
Trp Asp Met Met Met Asn Trp Ser Pro Thr Thr Ala Leu Val Val Ser	
1 5 10 15	
cag ctg ctc cgg atc cca caa gct gtc gtg gac atg gtg gcg ggg gcc	96
Gln Leu Leu Arg Ile Pro Gln Ala Val Val Asp Met Val Ala Gly Ala	
20 25 30	
cat tgg gga gtc ctg gcg ggc ctc gcc tac tat tcc atg gtg ggg aac	144
His Trp Gly Val Leu Ala Gly Leu Ala Tyr Tyr Ser Met Val Gly Asn	
35 40 45	
tgg gct aag gtt ttg gtt gtg atg cta ctc ttt gcc ggc gtc gac ggg	192
Trp Ala Lys Val Leu Val Val Met Leu Leu Phe Ala Gly Val Asp Gly	
50 55 60	
cat acc cgc gtg tca gga ggg gca gca gcc tcc gat acc agg ggc ctt	240
His Thr Arg Val Ser Gly Gly Ala Ala Ala Ser Asp Thr Arg Gly Leu	
65 70 75 80	
gtg tcc ctc ttt agc ccc ggg tcg gct cag aaa atc cag ctc gta aac	288
Val Ser Leu Phe Ser Pro Gly Ser Ala Gln Lys Ile Gln Leu Val Asn	
85 90 95	
acc aac ggc agt tgg cac atc aac agg act gcc ctg aac tgc aac gac	336
Thr Asn Gly Ser Trp His Ile Asn Arg Thr Ala Leu Asn Cys Asn Asp	
100 105 110	
tcc ctc caa aca ggg ttc ttt gcc gca cta ttc tac aaa cac aaa ttc	384
Ser Leu Gln Thr Gly Phe Phe Ala Ala Leu Phe Tyr Lys His Lys Phe	
115 120 125	
aac tcg tct gga tgc cca gag cgc ttg gcc agc tgt cgc tcc atc gac	432
Asn Ser Ser Gly Cys Pro Glu Arg Leu Ala Ser Cys Arg Ser Ile Asp	
130 135 140	
aag ttc gct cag ggg tgg ggt ccc ctc act tac act gag cct aac agc	480
Lys Phe Ala Gln Gly Trp Gly Pro Leu Thr Tyr Thr Glu Pro Asn Ser	
145 150 155 160	
tcg gac cag agg ccc tac tgc tgg cac tac gcg cct cga ccg tgt ggt	528
Ser Asp Gln Arg Pro Tyr Cys Trp His Tyr Ala Pro Arg Pro Cys Gly	
165 170 175	
att gta ccc gcg tct cag gtg tgc ggt cca gtg tat tgc ttc acc ccg	576
Ile Val Pro Ala Ser Gln Val Cys Gly Pro Val Tyr Cys Phe Thr Pro	
180 185 190	
agc cct gtt gtg gtg ggg acg acc gat cgg ttt ggt gtc ccc acg tat	624
Ser Pro Val Val Val Gly Thr Thr Asp Arg Phe Gly Val Pro Thr Tyr	
195 200 205	

aac tgg ggg gcg aac gac tcg gat gtg ctg att ctc aac aac acg cgg Asn Trp Gly Ala Asn Asp Ser Asp Val Leu Ile Leu Asn Asn Thr Arg 210 215 220	672
ccg ccg cga ggc aac tgg ttc ggc tgt aca tgg atg aat ggc act ggg Pro Pro Arg Gly Asn Trp Phe Gly Cys Thr Trp Met Asn Gly Thr Gly 225 230 235 240	720
ttc acc aag acg tgt ggg ggc ccc ccg tgc aac atc ggg ggg gcc ggc Phe Thr Lys Thr Cys Gly Gly Pro Pro Cys Asn Ile Gly Gly Ala Gly 245 250 255	768
aac aac acc ttg acc tgc ccc act gac tgt ttt cgg aag cac ccc gag Asn Asn Thr Leu Thr Cys Pro Thr Asp Cys Phe Arg Lys His Pro Glu 260 265 270	816
gcc acc tac gcc aga tgc ggt tct ggg ccc tgg ctg aca cct agg tgt Ala Thr Tyr Ala Arg Cys Gly Ser Gly Pro Trp Leu Thr Pro Arg Cys 275 280 285	864
atg gtt cat tac cca tat agg ctc tgg cac tac ccc tgc act gtc aac Met Val His Tyr Pro Tyr Arg Leu Trp His Tyr Pro Cys Thr Val Asn 290 295 300	912
ttc acc atc ttc aag gtt agg atg tac gtg ggg ggc gtg gag cac agg Phe Thr Ile Phe Lys Val Arg Met Tyr Val Gly Gly Val Glu His Arg 305 310 315 320	960
ttc gaa gcc gca tgc aat tgg act cga gga gag cgt tgt gac ttg gag Phe Glu Ala Ala Cys Asn Trp Thr Arg Gly Glu Arg Cys Asp Leu Glu 325 330 335	1008
gac agg gat aga tca gag ctt agc ccg ctg ctg ctg tct aca aca gag Asp Arg Asp Arg Ser Glu Leu Ser Pro Leu Leu Leu Ser Thr Thr Glu 340 345 350	1056
tgg cag ata ctg ccc tgt tcc ttc acc acc ctg ccg gcc cta tcc acc Trp Gln Ile Leu Pro Cys Ser Phe Thr Thr Leu Pro Ala Leu Ser Thr 355 360 365	1104
ggc ctg atc cac ctc cat cag aac atc gtg gac gtg caa tac ctg tac Gly Leu Ile His Leu His Gln Asn Ile Val Asp Val Gln Tyr Leu Tyr 370 375 380	1152
ggt gta ggg tcg gcg gtt gtc tcc ctt gtc atc aaa tgg gag tat gtc Gly Val Gly Ser Ala Val Val Ser Leu Val Ile Lys Trp Glu Tyr Val 385 390 395 400	1200
ctg ttg ctc ttc ctt ctc ctg gca gac gcg cgc atc tgc gcc tgc tta Leu Leu Leu Phe Leu Leu Leu Ala Asp Ala Arg Ile Cys Ala Cys Leu 405 410 415	1248
tgg atg atg ctg ctg ata gct caa gct gag gcc gcc tta gag aac ctg Trp Met Met Leu Leu Ile Ala Gln Ala Glu Ala Ala Leu Glu Asn Leu 420 425 430	1296
gtg gtc ctc aat gcg gcg gcc gtg gcc ggg gcg cat ggc act ctt tcc Val Val Leu Asn Ala Ala Ala Val Ala Gly Ala His Gly Thr Leu Ser 435 440 445	1344
ttc ctt gtg ttc ttc tgt gct gcc tgg tac atc aag ggc agg ctg gtc	1392

Phe Leu Val Phe Phe Cys Ala Ala Trp Tyr Ile Lys Gly Arg Leu Val  
 450 455 460

cct ggt gcg gca tac gcc ttc tat ggc gtg tgg ccg ctg ctc ctg ctt 1440  
 Pro Gly Ala Ala Tyr Ala Phe Tyr Gly Val Trp Pro Leu Leu Leu Leu  
 465 470 475 480

ctg ctg gcc tta cca cca cga gct tat gcc tagtaa 1476  
 Leu Leu Ala Leu Pro Pro Arg Ala Tyr Ala  
 485 490

<210> 36

<211> 490

<212> PRT

<213> Hepatitis C virus

<400> 36

Trp Asp Met Met Met Asn Trp Ser Pro Thr Thr Ala Leu Val Val Ser  
 1 5 10 15

Gln Leu Leu Arg Ile Pro Gln Ala Val Val Asp Met Val Ala Gly Ala  
 20 25 30

His Trp Gly Val Leu Ala Gly Leu Ala Tyr Tyr Ser Met Val Gly Asn  
 35 40 45

Trp Ala Lys Val Leu Val Val Met Leu Leu Phe Ala Gly Val Asp Gly  
 50 55 60

His Thr Arg Val Ser Gly Gly Ala Ala Ala Ser Asp Thr Arg Gly Leu  
 65 70 75 80

Val Ser Leu Phe Ser Pro Gly Ser Ala Gln Lys Ile Gln Leu Val Asn  
 85 90 95

Thr Asn Gly Ser Trp His Ile Asn Arg Thr Ala Leu Asn Cys Asn Asp  
 100 105 110

Ser Leu Gln Thr Gly Phe Phe Ala Ala Leu Phe Tyr Lys His Lys Phe  
 115 120 125

Asn Ser Ser Gly Cys Pro Glu Arg Leu Ala Ser Cys Arg Ser Ile Asp  
 130 135 140

Lys Phe Ala Gln Gly Trp Gly Pro Leu Thr Tyr Thr Glu Pro Asn Ser  
 145 150 155 160

Ser Asp Gln Arg Pro Tyr Cys Trp His Tyr Ala Pro Arg Pro Cys Gly  
 165 170 175

Ile Val Pro Ala Ser Gln Val Cys Gly Pro Val Tyr Cys Phe Thr Pro  
 180 185 190

Ser Pro Val Val Val Gly Thr Thr Asp Arg Phe Gly Val Pro Thr Tyr  
 195 200 205

Asn Trp Gly Ala Asn Asp Ser Asp Val Leu Ile Leu Asn Asn Thr Arg  
 210 215 220

Pro Pro Arg Gly Asn Trp Phe Gly Cys Thr Trp Met Asn Gly Thr Gly



225		230		235		240
Phe Thr Lys Thr Cys Gly Gly Pro Pro Cys Asn Ile Gly Gly Ala Gly						
	245			250		255
Asn Asn Thr Leu Thr Cys Pro Thr Asp Cys Phe Arg Lys His Pro Glu						
	260		265			270
Ala Thr Tyr Ala Arg Cys Gly Ser Gly Pro Trp Leu Thr Pro Arg Cys						
	275		280			285
Met Val His Tyr Pro Tyr Arg Leu Trp His Tyr Pro Cys Thr Val Asn						
	290		295			300
Phe Thr Ile Phe Lys Val Arg Met Tyr Val Gly Gly Val Glu His Arg						
305		310		315		320
Phe Glu Ala Ala Cys Asn Trp Thr Arg Gly Glu Arg Cys Asp Leu Glu						
	325		330			335
Asp Arg Asp Arg Ser Glu Leu Ser Pro Leu Leu Leu Ser Thr Thr Glu						
	340		345			350
Trp Gln Ile Leu Pro Cys Ser Phe Thr Thr Leu Pro Ala Leu Ser Thr						
	355		360			365
Gly Leu Ile His Leu His Gln Asn Ile Val Asp Val Gln Tyr Leu Tyr						
	370		375			380
Gly Val Gly Ser Ala Val Val Ser Leu Val Ile Lys Trp Glu Tyr Val						
385		390		395		400
Leu Leu Leu Phe Leu Leu Leu Ala Asp Ala Arg Ile Cys Ala Cys Leu						
	405		410			415
Trp Met Met Leu Leu Ile Ala Gln Ala Glu Ala Ala Leu Glu Asn Leu						
	420		425			430
Val Val Leu Asn Ala Ala Ala Val Ala Gly Ala His Gly Thr Leu Ser						
	435		440			445
Phe Leu Val Phe Phe Cys Ala Ala Trp Tyr Ile Lys Gly Arg Leu Val						
	450		455			460
Pro Gly Ala Ala Tyr Ala Phe Tyr Gly Val Trp Pro Leu Leu Leu Leu						
465		470		475		480
Leu Leu Ala Leu Pro Pro Arg Ala Tyr Ala						
	485			490		

&lt;210&gt; 37

&lt;211&gt; 1021

&lt;212&gt; DNA

&lt;213&gt; Hepatitis C virus

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; 2..1018

&lt;220&gt;

&lt;221&gt; mat\_peptide

&lt;222&gt; 2..1015

&lt;400&gt; 37

g atc cca caa gct gtc gtg gac atg gtg gcg ggg gcc cat tgg gga 46  
 Ile Pro Gln Ala Val Val Asp Met Val Ala Gly Ala His Trp Gly  
 1 5 10 15

gtc ctg gcg ggc ctc gcc tac tat tcc atg gtg ggg aac tgg gct aag 94  
 Val Leu Ala Gly Leu Ala Tyr Tyr Ser Met Val Gly Asn Trp Ala Lys  
 20 25 30

gtt ttg gtt gtg atg cta ctc ttt gcc ggc gtc gac ggg cat acc cgc 142  
 Val Leu Val Val Met Leu Leu Phe Ala Gly Val Asp Gly His Thr Arg  
 35 40 45

gtg tca gga ggg gca gca gcc tcc gat acc agg ggc ctt gtg tcc ctc 190  
 Val Ser Gly Gly Ala Ala Ala Ser Asp Thr Arg Gly Leu Val Ser Leu  
 50 55 60

ttt agc ccc ggg tcg gct cag aaa atc cag ctc gta aac acc aac ggc 238  
 Phe Ser Pro Gly Ser Ala Gln Lys Ile Gln Leu Val Asn Thr Asn Gly  
 65 70 75

agt tgg cac atc aac agg act gcc ctg aac tgc aac gac tcc ctc caa 286  
 Ser Trp His Ile Asn Arg Thr Ala Leu Asn Cys Asn Asp Ser Leu Gln  
 80 85 90 95

aca ggg ttc ttt gcc gca cta ttc tac aaa cac aaa ttc aac tcg tct 334  
 Thr Gly Phe Phe Ala Ala Leu Phe Tyr Lys His Lys Phe Asn Ser Ser  
 100 105 110

gga tgc cca gag cgc ttg gcc agc tgt cgc tcc atc gac aag ttc gct 382  
 Gly Cys Pro Glu Arg Leu Ala Ser Cys Arg Ser Ile Asp Lys Phe Ala  
 115 120 125

cag ggg tgg ggt ccc ctc act tac act gag cct aac agc tcg gac cag 430  
 Gln Gly Trp Gly Pro Leu Thr Tyr Thr Glu Pro Asn Ser Ser Asp Gln  
 130 135 140

agg ccc tac tgc tgg cac tac gcg cct cga ccg tgt ggt att gta ccc 478  
 Arg Pro Tyr Cys Trp His Tyr Ala Pro Arg Pro Cys Gly Ile Val Pro  
 145 150 155

gcg tct cag gtg tgc ggt cca gtg tat tgc ttc acc ccg agc cct gtt 526  
 Ala Ser Gln Val Cys Gly Pro Val Tyr Cys Phe Thr Pro Ser Pro Val  
 160 165 170 175

gtg gtg ggg acg acc gat cgg ttt ggt gtc ccc acg tat aac tgg ggg 574  
 Val Val Gly Thr Thr Asp Arg Phe Gly Val Pro Thr Tyr Asn Trp Gly  
 180 185 190

gcg aac gac tcg gat gtg ctg att ctc aac aac acg cgg ccg ccg cga 622  
 Ala Asn Asp Ser Asp Val Leu Ile Leu Asn Asn Thr Arg Pro Pro Arg  
 195 200 205

ggc aac tgg ttc ggc tgt aca tgg atg aat ggc act ggg ttc acc aag 670  
 Gly Asn Trp Phe Gly Cys Thr Trp Met Asn Gly Thr Gly Phe Thr Lys  
 210 215 220

acg tgt ggg ggc ccc ccg tgc aac atc ggg ggg gcc ggc aac aac acc 718

Thr Cys Gly Gly Pro Pro Cys Asn Ile Gly Gly Ala Gly Asn Asn Thr  
 225 230 235  
 ttg acc tgc ccc act gac tgt ttt cgg aag cac ccc gag gcc acc tac 766  
 Leu Thr Cys Pro Thr Asp Cys Phe Arg Lys His Pro Glu Ala Thr Tyr  
 240 245 250 255  
 gcc aga tgc ggt tct ggg ccc tgg ctg aca cct agg tgt atg gtt cat 814  
 Ala Arg Cys Gly Ser Gly Pro Trp Leu Thr Pro Arg Cys Met Val His  
 260 265 270  
 tac cca tat agg ctc tgg cac tac ccc tgc act gtc aac ttc acc atc 862  
 Tyr Pro Tyr Arg Leu Trp His Tyr Pro Cys Thr Val Asn Phe Thr Ile  
 275 280 285  
 ttc aag gtt agg atg tac gtg ggg ggc gtg gag cac agg ttc gaa gcc 910  
 Phe Lys Val Arg Met Tyr Val Gly Gly Val Glu His Arg Phe Glu Ala  
 290 295 300  
 gca tgc aat tgg act cga gga gag cgt tgt gac ttg gag gac agg gat 958  
 Ala Cys Asn Trp Thr Arg Gly Glu Arg Cys Asp Leu Glu Asp Arg Asp  
 305 310 315  
 aga tca gag ctt agc ccg ctg ctg ctg tct aca aca gag tgg cag agt 1006  
 Arg Ser Glu Leu Ser Pro Leu Leu Leu Ser Thr Thr Glu Trp Gln Ser  
 320 325 330 335  
 ggc aga gct taatta 1021  
 Gly Arg Ala  
 <210> 38  
 <211> 338  
 <212> PRT  
 <213> Hepatitis C virus  
 <400> 38  
 Ile Pro Gln Ala Val Val Asp Met Val Ala Gly Ala His Trp Gly Val  
 1 5 10 15  
 Leu Ala Gly Leu Ala Tyr Tyr Ser Met Val Gly Asn Trp Ala Lys Val  
 20 25 30  
 Leu Val Val Met Leu Leu Phe Ala Gly Val Asp Gly His Thr Arg Val  
 35 40 45  
 Ser Gly Gly Ala Ala Ala Ser Asp Thr Arg Gly Leu Val Ser Leu Phe  
 50 55 60  
 Ser Pro Gly Ser Ala Gln Lys Ile Gln Leu Val Asn Thr Asn Gly Ser  
 65 70 75 80  
 Trp His Ile Asn Arg Thr Ala Leu Asn Cys Asn Asp Ser Leu Gln Thr  
 85 90 95  
 Gly Phe Phe Ala Ala Leu Phe Tyr Lys His Lys Phe Asn Ser Ser Gly  
 100 105 110  
 Cys Pro Glu Arg Leu Ala Ser Cys Arg Ser Ile Asp Lys Phe Ala Gln

115                      120                      125  
 Gly Trp Gly Pro Leu Thr Tyr Thr Glu Pro Asn Ser Ser Asp Gln Arg  
 130                      135                      140  
 Pro Tyr Cys Trp His Tyr Ala Pro Arg Pro Cys Gly Ile Val Pro Ala  
 145                      150                      155                      160  
 Ser Gln Val Cys Gly Pro Val Tyr Cys Phe Thr Pro Ser Pro Val Val  
 165                      170                      175  
 Val Gly Thr Thr Asp Arg Phe Gly Val Pro Thr Tyr Asn Trp Gly Ala  
 180                      185                      190  
 Asn Asp Ser Asp Val Leu Ile Leu Asn Asn Thr Arg Pro Pro Arg Gly  
 195                      200                      205  
 Asn Trp Phe Gly Cys Thr Trp Met Asn Gly Thr Gly Phe Thr Lys Thr  
 210                      215                      220  
 Cys Gly Gly Pro Pro Cys Asn Ile Gly Gly Ala Gly Asn Asn Thr Leu  
 225                      230                      235                      240  
 Thr Cys Pro Thr Asp Cys Phe Arg Lys His Pro Glu Ala Thr Tyr Ala  
 245                      250                      255  
 Arg Cys Gly Ser Gly Pro Trp Leu Thr Pro Arg Cys Met Val His Tyr  
 260                      265                      270  
 Pro Tyr Arg Leu Trp His Tyr Pro Cys Thr Val Asn Phe Thr Ile Phe  
 275                      280                      285  
 Lys Val Arg Met Tyr Val Gly Gly Val Glu His Arg Phe Glu Ala Ala  
 290                      295                      300  
 Cys Asn Trp Thr Arg Gly Glu Arg Cys Asp Leu Glu Asp Arg Asp Arg  
 305                      310                      315                      320  
 Ser Glu Leu Ser Pro Leu Leu Leu Ser Thr Thr Glu Trp Gln Ser Gly  
 325                      330                      335

Arg Ala

<210> 39

<211> 1034

<212> DNA

<213> Hepatitis C virus

<220>

<221> CDS

<222> 2..1032

<220>

<221> mat\_peptide

<222> 2..1029

<400> 39

g atc cca caa gct gtc gtg gac atg gtg gcg ggg gcc cat tgg gga

Ile	Pro	Gln	Ala	Val	Val	Asp	Met	Val	Ala	Gly	Ala	His	Trp	Gly		
1				5					10					15		
gtc	ctg	gcg	ggc	ctc	gcc	tac	tat	tcc	atg	gtg	ggg	aac	tgg	gct	aag	94
Val	Leu	Ala	Gly	Leu	Ala	Tyr	Tyr	Ser	Met	Val	Gly	Asn	Trp	Ala	Lys	
				20				25					30			
ggt	ttg	ggt	gtg	atg	cta	ctc	ttt	gcc	ggc	gtc	gac	ggg	cat	acc	cgc	142
Val	Leu	Val	Val	Met	Leu	Leu	Phe	Ala	Gly	Val	Asp	Gly	His	Thr	Arg	
				35				40					45			
gtg	tca	gga	ggg	gca	gca	gcc	tcc	gat	acc	agg	ggc	ctt	gtg	tcc	ctc	190
Val	Ser	Gly	Gly	Ala	Ala	Ala	Ser	Asp	Thr	Arg	Gly	Leu	Val	Ser	Leu	
		50					55					60				
ttt	agc	ccc	ggg	tcg	gct	cag	aaa	atc	cag	ctc	gta	aac	acc	aac	ggc	238
Phe	Ser	Pro	Gly	Ser	Ala	Gln	Lys	Ile	Gln	Leu	Val	Asn	Thr	Asn	Gly	
	65					70				75						
agt	tgg	cac	atc	aac	agg	act	gcc	ctg	aac	tgc	aac	gac	tcc	ctc	caa	286
Ser	Trp	His	Ile	Asn	Arg	Thr	Ala	Leu	Asn	Cys	Asn	Asp	Ser	Leu	Gln	
	80				85				90					95		
aca	ggg	ttc	ttt	gcc	gca	cta	ttc	tac	aaa	cac	aaa	ttc	aac	tcg	tct	334
Thr	Gly	Phe	Phe	Ala	Ala	Leu	Phe	Tyr	Lys	His	Lys	Phe	Asn	Ser	Ser	
				100					105					110		
gga	tgc	cca	gag	cgc	ttg	gcc	agc	tgt	cgc	tcc	atc	gac	aag	ttc	gct	382
Gly	Cys	Pro	Glu	Arg	Leu	Ala	Ser	Cys	Arg	Ser	Ile	Asp	Lys	Phe	Ala	
			115					120					125			
cag	ggg	tgg	ggt	ccc	ctc	act	tac	act	gag	cct	aac	agc	tcg	gac	cag	430
Gln	Gly	Trp	Gly	Pro	Leu	Thr	Tyr	Thr	Glu	Pro	Asn	Ser	Ser	Asp	Gln	
		130					135					140				
agg	ccc	tac	tgc	tgg	cac	tac	gcg	cct	cga	ccg	tgt	ggt	att	gta	ccc	478
Arg	Pro	Tyr	Cys	Trp	His	Tyr	Ala	Pro	Arg	Pro	Cys	Gly	Ile	Val	Pro	
	145					150					155					
gcg	tct	cag	gtg	tgc	ggt	cca	gtg	tat	tgc	ttc	acc	ccg	agc	cct	ggt	526
Ala	Ser	Gln	Val	Cys	Gly	Pro	Val	Tyr	Cys	Phe	Thr	Pro	Ser	Pro	Val	
	160				165				170					175		
gtg	gtg	ggg	acg	acc	gat	cgg	ttt	ggt	gtc	ccc	acg	tat	aac	tgg	ggg	574
Val	Val	Gly	Thr	Thr	Asp	Arg	Phe	Gly	Val	Pro	Thr	Tyr	Asn	Trp	Gly	
				180					185					190		
gcg	aac	gac	tcg	gat	gtg	ctg	att	ctc	aac	aac	acg	cgg	ccg	ccg	cga	622
Ala	Asn	Asp	Ser	Asp	Val	Leu	Ile	Leu	Asn	Asn	Thr	Arg	Pro	Pro	Arg	
			195					200					205			
ggc	aac	tgg	ttc	ggc	tgt	aca	tgg	atg	aat	ggc	act	ggg	ttc	acc	aag	670
Gly	Asn	Trp	Phe	Gly	Cys	Thr	Trp	Met	Asn	Gly	Thr	Gly	Phe	Thr	Lys	
		210					215					220				
acg	tgt	ggg	ggc	ccc	ccg	tgc	aac	atc	ggg	ggg	gcc	ggc	aac	aac	acc	718
Thr	Cys	Gly	Gly	Pro	Pro	Cys	Asn	Ile	Gly	Gly	Ala	Gly	Asn	Asn	Thr	
	225					230					235					
ttg	acc	tgc	ccc	act	gac	tgt	ttt	cgg	aag	cac	ccc	gag	gcc	acc	tac	766
Leu	Thr	Cys	Pro	Thr	Asp	Cys	Phe	Arg	Lys	His	Pro	Glu	Ala	Thr	Tyr	
	240				245					250				255		

gcc aga tgc ggt tct ggg ccc tgg ctg aca cct agg tgt atg gtt cat 814  
Ala Arg Cys Gly Ser Gly Pro Trp Leu Thr Pro Arg Cys Met Val His  
260 265 270

tac cca tat agg ctc tgg cac tac ccc tgc act gtc aac ttc acc atc 862  
Tyr Pro Tyr Arg Leu Trp His Tyr Pro Cys Thr Val Asn Phe Thr Ile  
275 280 285

ttc aag gtt agg atg tac gtg ggg ggc gtg gag cac agg ttc gaa gcc 910  
Phe Lys Val Arg Met Tyr Val Gly Gly Val Glu His Arg Phe Glu Ala  
290 295 300

gca tgc aat tgg act cga gga gag cgt tgt gac ttg gag gac agg gat 958  
Ala Cys Asn Trp Thr Arg Gly Glu Arg Cys Asp Leu Glu Asp Arg Asp  
305 310 315

aga tca gag ctt agc ccg ctg ctg ctg tct aca aca ggt gat cga ggg 1006  
Arg Ser Glu Leu Ser Pro Leu Leu Leu Ser Thr Thr Gly Asp Arg Gly  
320 325 330 335

cag aca cca tca cca cca tca cta at ag 1034  
Gln Thr Pro Ser Pro Pro Ser Leu  
340

<210> 40

<211> 343

<212> PRT

<213> Hepatitis C virus

<400> 40

Ile Pro Gln Ala Val Val Asp Met Val Ala Gly Ala His Trp Gly Val  
1 5 10 15

Leu Ala Gly Leu Ala Tyr Tyr Ser Met Val Gly Asn Trp Ala Lys Val  
20 25 30

Leu Val Val Met Leu Leu Phe Ala Gly Val Asp Gly His Thr Arg Val  
35 40 45

Ser Gly Gly Ala Ala Ala Ser Asp Thr Arg Gly Leu Val Ser Leu Phe  
50 55 60

Ser Pro Gly Ser Ala Gln Lys Ile Gln Leu Val Asn Thr Asn Gly Ser  
65 70 75 80

Trp His Ile Asn Arg Thr Ala Leu Asn Cys Asn Asp Ser Leu Gln Thr  
85 90 95

Gly Phe Phe Ala Ala Leu Phe Tyr Lys His Lys Phe Asn Ser Ser Gly  
100 105 110

Cys Pro Glu Arg Leu Ala Ser Cys Arg Ser Ile Asp Lys Phe Ala Gln  
115 120 125

Gly Trp Gly Pro Leu Thr Tyr Thr Glu Pro Asn Ser Ser Asp Gln Arg  
130 135 140

Pro Tyr Cys Trp His Tyr Ala Pro Arg Pro Cys Gly Ile Val Pro Ala

145                      150                      155                      160  
 Ser Gln Val Cys Gly Pro Val Tyr Cys Phe Thr Pro Ser Pro Val Val  
                                  165                      170                      175  
 Val Gly Thr Thr Asp Arg Phe Gly Val Pro Thr Tyr Asn Trp Gly Ala  
                                  180                      185                      190  
 Asn Asp Ser Asp Val Leu Ile Leu Asn Asn Thr Arg Pro Pro Arg Gly  
                                  195                      200                      205  
 Asn Trp Phe Gly Cys Thr Trp Met Asn Gly Thr Gly Phe Thr Lys Thr  
                                  210                      215                      220  
 Cys Gly Gly Pro Pro Cys Asn Ile Gly Gly Ala Gly Asn Asn Thr Leu  
                                  225                      230                      235                      240  
 Thr Cys Pro Thr Asp Cys Phe Arg Lys His Pro Glu Ala Thr Tyr Ala  
                                  245                      250                      255  
 Arg Cys Gly Ser Gly Pro Trp Leu Thr Pro Arg Cys Met Val His Tyr  
                                  260                      265                      270  
 Pro Tyr Arg Leu Trp His Tyr Pro Cys Thr Val Asn Phe Thr Ile Phe  
                                  275                      280                      285  
 Lys Val Arg Met Tyr Val Gly Gly Val Glu His Arg Phe Glu Ala Ala  
                                  290                      295                      300  
 Cys Asn Trp Thr Arg Gly Glu Arg Cys Asp Leu Glu Asp Arg Asp Arg  
                                  305                      310                      315                      320  
 Ser Glu Leu Ser Pro Leu Leu Leu Ser Thr Thr Gly Asp Arg Gly Gln  
                                  325                      330                      335  
 Thr Pro Ser Pro Pro Ser Leu  
                                  340

&lt;210&gt; 41

&lt;211&gt; 945

&lt;212&gt; DNA

&lt;213&gt; Hepatitis C virus

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; 1..942

&lt;220&gt;

&lt;221&gt; mat\_peptide

&lt;222&gt; 1..939

&lt;400&gt; 41

atg gtg ggg aac tgg gct aag gtt ttg gtt gtg atg cta ctc ttt gcc                      48  
 Met Val Gly Asn Trp Ala Lys Val Leu Val Val Met Leu Leu Phe Ala  
   1                                      5                                      10                                      15  
 ggc gtc gac ggg cat acc cgc gtg tca gga ggg gca gca gcc tcc gat                      96  
 Gly Val Asp Gly His Thr Arg Val Ser Gly Gly Ala Ala Ala Ser Asp  
                                  20                                      25                                      30  
 acc agg ggc ctt gtg tcc ctc ttt agc ccc ggg tcg gct cag aaa atc                      144  
 Thr Arg Gly Leu Val Ser Leu Phe Ser Pro Gly Ser Ala Gln Lys Ile

35	40	45	
cag. ctc gta aac acc aac ggc agt tgg cac atc aac agg act gcc ctg			192
Gln Leu Val Asn Thr Asn Gly Ser Trp His Ile Asn Arg Thr Ala Leu			
50	55	60	
aac tgc aac gac tcc ctc caa aca ggg ttc ttt gcc gca cta ttc tac			240
Asn Cys Asn Asp Ser Leu Gln Thr Gly Phe Phe Ala Ala Leu Phe Tyr			
65	70	75	80
aaa cac aaa ttc aac tcg tct gga tgc cca gag cgc ttg gcc agc tgt			288
Lys His Lys Phe Asn Ser Ser Gly Cys Pro Glu Arg Leu Ala Ser Cys			
	85	90	95
cgc tcc atc gac aag ttc gct cag ggg tgg ggt ccc ctc act tac act			336
Arg Ser Ile Asp Lys Phe Ala Gln Gly Trp Gly Pro Leu Thr Tyr Thr			
	100	105	110
gag cct aac agc tcg gac cag agg ccc tac tgc tgg cac tac gcg cct			384
Glu Pro Asn Ser Ser Asp Gln Arg Pro Tyr Cys Trp His Tyr Ala Pro			
	115	120	125
cga ccg tgt ggt att gta ccc gcg tct cag gtg tgc ggt cca gtg tat			432
Arg Pro Cys Gly Ile Val Pro Ala Ser Gln Val Cys Gly Pro Val Tyr			
	130	135	140
tgc ttc acc ccg agc cct gtt gtg gtg ggg acg acc gat cgg ttt ggt			480
Cys Phe Thr Pro Ser Pro Val Val Val Gly Thr Thr Asp Arg Phe Gly			
145	150	155	160
gtc ccc acg tat aac tgg ggg gcg aac gac tcg gat gtg ctg att ctc			528
Val Pro Thr Tyr Asn Trp Gly Ala Asn Asp Ser Asp Val Leu Ile Leu			
	165	170	175
aac aac acg ccg ccg ccg cga ggc aac tgg ttc ggc tgt aca tgg atg			576
Asn Asn Thr Arg Pro Pro Arg Gly Asn Trp Phe Gly Cys Thr Trp Met			
	180	185	190
aat ggc act ggg ttc acc aag acg tgt ggg ggc ccc ccg tgc aac atc			624
Asn Gly Thr Gly Phe Thr Lys Thr Cys Gly Gly Pro Pro Cys Asn Ile			
	195	200	205
ggg ggg gcc gcc aac aac acc ttg acc tgc ccc act gac tgt ttt cgg			672
Gly Gly Ala Gly Asn Asn Thr Leu Thr Cys Pro Thr Asp Cys Phe Arg			
	210	215	220
aag cac ccc gag gcc acc tac gcc aga tgc ggt tct ggg ccc tgg ctg			720
Lys His Pro Glu Ala Thr Tyr Ala Arg Cys Gly Ser Gly Pro Trp Leu			
225	230	235	240
aca cct agg tgt atg gtt cat tac cca tat agg ctc tgg cac tac ccc			768
Thr Pro Arg Cys Met Val His Tyr Pro Tyr Arg Leu Trp His Tyr Pro			
	245	250	255
tgc act gtc aac ttc acc atc ttc aag gtt agg atg tac gtg ggg ggc			816
Cys Thr Val Asn Phe Thr Ile Phe Lys Val Arg Met Tyr Val Gly Gly			
	260	265	270
gtg gag cac agg ttc gaa gcc gca tgc aat tgg act cga gga gag cgt			864
Val Glu His Arg Phe Glu Ala Ala Cys Asn Trp Thr Arg Gly Glu Arg			
	275	280	285



tgt gac ttg gag gac agg gat aga tca gag ctt agc ccg ctg ctg ctg 912  
 Cys Asp Leu Glu Asp Arg Asp Arg Ser Glu Leu Ser Pro Leu Leu Leu  
 290 295 300

tct aca aca gag tgg cag agc tta att aat tag 945  
 Ser Thr Thr Glu Trp Gln Ser Leu Ile Asn  
 305 310

<210> 42

<211> 314

<212> PRT

<213> Hepatitis C virus

<400> 42

Met Val Gly Asn Trp Ala Lys Val Leu Val Val Met Leu Leu Phe Ala  
 1 5 10 15

Gly Val Asp Gly His Thr Arg Val Ser Gly Gly Ala Ala Ala Ser Asp  
 20 25 30

Thr Arg Gly Leu Val Ser Leu Phe Ser Pro Gly Ser Ala Gln Lys Ile  
 35 40 45

Gln Leu Val Asn Thr Asn Gly Ser Trp His Ile Asn Arg Thr Ala Leu  
 50 55 60

Asn Cys Asn Asp Ser Leu Gln Thr Gly Phe Phe Ala Ala Leu Phe Tyr  
 65 70 75 80

Lys His Lys Phe Asn Ser Ser Gly Cys Pro Glu Arg Leu Ala Ser Cys  
 85 90 95

Arg Ser Ile Asp Lys Phe Ala Gln Gly Trp Gly Pro Leu Thr Tyr Thr  
 100 105 110

Glu Pro Asn Ser Ser Asp Gln Arg Pro Tyr Cys Trp His Tyr Ala Pro  
 115 120 125

Arg Pro Cys Gly Ile Val Pro Ala Ser Gln Val Cys Gly Pro Val Tyr  
 130 135 140

Cys Phe Thr Pro Ser Pro Val Val Val Gly Thr Thr Asp Arg Phe Gly  
 145 150 155 160

Val Pro Thr Tyr Asn Trp Gly Ala Asn Asp Ser Asp Val Leu Ile Leu  
 165 170 175

Asn Asn Thr Arg Pro Pro Arg Gly Asn Trp Phe Gly Cys Thr Trp Met  
 180 185 190

Asn Gly Thr Gly Phe Thr Lys Thr Cys Gly Gly Pro Pro Cys Asn Ile  
 195 200 205

Gly Gly Ala Gly Asn Asn Thr Leu Thr Cys Pro Thr Asp Cys Phe Arg  
 210 215 220

Lys His Pro Glu Ala Thr Tyr Ala Arg Cys Gly Ser Gly Pro Trp Leu  
 225 230 235 240

Thr Pro Arg Cys Met Val His Tyr Pro Tyr Arg Leu Trp His Tyr Pro  
 245 250 255

Cys Thr Val Asn Phe Thr Ile Phe Lys Val Arg Met Tyr Val Gly Gly  
 260 265 270

Val Glu His Arg Phe Glu Ala Ala Cys Asn Trp Thr Arg Gly Glu Arg  
 275 280 285

Cys Asp Leu Glu Asp Arg Asp Arg Ser Glu Leu Ser Pro Leu Leu Leu  
 290 295 300

Ser Thr Thr Glu Trp Gln Ser Leu Ile Asn  
 305 310

&lt;210&gt; 43

&lt;211&gt; 961

&lt;212&gt; DNA

&lt;213&gt; Hepatitis C virus

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; 1..958

&lt;220&gt;

&lt;221&gt; mat\_peptide

&lt;222&gt; 1..955

&lt;400&gt; 43

atg gtg ggg aac tgg gct aag gtt ttg gtt gtg atg cta ctc ttt gcc	48
Met Val Gly Asn Trp Ala Lys Val Leu Val Val Met Leu Leu Phe Ala	
1 5 10 15	
ggc gtc gac ggg cat acc cgc gtg tca gga ggg gca gca gcc tcc gat	96
Gly Val Asp Gly His Thr Arg Val Ser Gly Gly Ala Ala Ser Asp	
20 25 30	
acc agg ggc ctt gtg tcc ctc ttt agc ccc ggg tcg gct cag aaa atc	144
Thr Arg Gly Leu Val Ser Leu Phe Ser Pro Gly Ser Ala Gln Lys Ile	
35 40 45	
cag ctc gta aac acc aac ggc agt tgg cac atc aac agg act gcc ctg	192
Gln Leu Val Asn Thr Asn Gly Ser Trp His Ile Asn Arg Thr Ala Leu	
50 55 60	
aac tgc aac gac tcc ctc caa aca ggg ttc ttt gcc gca cta ttc tac	240
Asn Cys Asn Asp Ser Leu Gln Thr Gly Phe Phe Ala Ala Leu Phe Tyr	
65 70 75 80	
aaa cac aaa ttc aac tcg tct gga tgc cca gag cgc ttg gcc agc tgt	288
Lys His Lys Phe Asn Ser Ser Gly Cys Pro Glu Arg Leu Ala Ser Cys	
85 90 95	
cgc tcc atc gac aag ttc gct cag ggg tgg ggt ccc ctc act tac act	336
Arg Ser Ile Asp Lys Phe Ala Gln Gly Trp Gly Pro Leu Thr Tyr Thr	
100 105 110	
gag cct aac agc tcg gac cag agg ccc tac tgc tgg cac tac gcg cct	384
Glu Pro Asn Ser Ser Asp Gln Arg Pro Tyr Cys Trp His Tyr Ala Pro	
115 120 125	

cga ccg tgt ggt att gta ccc gcg tct cag gtg tgc ggt cca gtg tat 432  
 Arg Pro Cys Gly Ile Val Pro Ala Ser Gln Val Cys Gly Pro Val Tyr  
 130 135 140

tgc ttc acc ccg agc cct gtt gtg gtg ggg acg acc gat cgg ttt ggt 480  
 Cys Phe Thr Pro Ser Pro Val Val Val Gly Thr Thr Asp Arg Phe Gly  
 145 150 155 160

gtc ccc acg tat aac tgg ggg gcg aac gac tcg gat gtg ctg att ctc 528  
 Val Pro Thr Tyr Asn Trp Gly Ala Asn Asp Ser Asp Val Leu Ile Leu  
 165 170 175

aac aac acg ccg ccg ccg cga ggc aac tgg ttc ggc tgt aca tgg atg 576  
 Asn Asn Thr Arg Pro Pro Arg Gly Asn Trp Phe Gly Cys Thr Trp Met  
 180 185 190

aat ggc act ggg ttc acc aag acg tgt ggg ggc ccc ccg tgc aac atc 624  
 Asn Gly Thr Gly Phe Thr Lys Thr Cys Gly Gly Pro Pro Cys Asn Ile  
 195 200 205

ggg ggg gcc ggc aac aac acc ttg acc tgc ccc act gac tgt ttt cgg 672  
 Gly Gly Ala Gly Asn Asn Thr Leu Thr Cys Pro Thr Asp Cys Phe Arg  
 210 215 220

aag cac ccc gag gcc acc tac gcc aga tgc ggt tct ggg ccc tgg ctg 720  
 Lys His Pro Glu Ala Thr Tyr Ala Arg Cys Gly Ser Gly Pro Trp Leu  
 225 230 235 240

aca cct agg tgt atg gtt cat tac cca tat agg ctc tgg cac tac ccc 768  
 Thr Pro Arg Cys Met Val His Tyr Pro Tyr Arg Leu Trp His Tyr Pro  
 245 250 255

tgc act gtc aac ttc acc atc ttc aag gtt agg atg tac gtg ggg ggc 816  
 Cys Thr Val Asn Phe Thr Ile Phe Lys Val Arg Met Tyr Val Gly Gly  
 260 265 270

gtg gag cac agg ttc gaa gcc gca tgc aat tgg act cga gga gag cgt 864  
 Val Glu His Arg Phe Glu Ala Ala Cys Asn Trp Thr Arg Gly Glu Arg  
 275 280 285

tgt gac ttg gag gac agg gat aga tca gag ctt agc ccg ctg ctg ctg 912  
 Cys Asp Leu Glu Asp Arg Asp Arg Ser Glu Leu Ser Pro Leu Leu Leu  
 290 295 300

tct aca aca ggt gat cga ggg cag aca cca tca cca cca tca cta a 958  
 Ser Thr Thr Gly Asp Arg Gly Gln Thr Pro Ser Pro Pro Ser Leu  
 305 310 315

tag 961  
 <210> 44  
 <211> 319  
 <212> PRT  
 <213> Hepatitis C virus  
 <400> 44  
 Met Val Gly Asn Trp Ala Lys Val Leu Val Val Met Leu Leu Phe Ala  
 1 5 10 15

Gly Val Asp Gly His Thr Arg Val Ser Gly Gly Ala Ala Ala Ser Asp  
 20 25 30  
 Thr Arg Gly Leu Val Ser Leu Phe Ser Pro Gly Ser Ala Gln Lys Ile  
 35 40 45  
 Gln Leu Val Asn Thr Asn Gly Ser Trp His Ile Asn Arg Thr Ala Leu  
 50 55 60  
 Asn Cys Asn Asp Ser Leu Gln Thr Gly Phe Phe Ala Ala Leu Phe Tyr  
 65 70 75 80  
 Lys His Lys Phe Asn Ser Ser Gly Cys Pro Glu Arg Leu Ala Ser Cys  
 85 90 95  
 Arg Ser Ile Asp Lys Phe Ala Gln Gly Trp Gly Pro Leu Thr Tyr Thr  
 100 105 110  
 Glu Pro Asn Ser Ser Asp Gln Arg Pro Tyr Cys Trp His Tyr Ala Pro  
 115 120 125  
 Arg Pro Cys Gly Ile Val Pro Ala Ser Gln Val Cys Gly Pro Val Tyr  
 130 135 140  
 Cys Phe Thr Pro Ser Pro Val Val Val Gly Thr Thr Asp Arg Phe Gly  
 145 150 155 160  
 Val Pro Thr Tyr Asn Trp Gly Ala Asn Asp Ser Asp Val Leu Ile Leu  
 165 170 175  
 Asn Asn Thr Arg Pro Pro Arg Gly Asn Trp Phe Gly Cys Thr Trp Met  
 180 185 190  
 Asn Gly Thr Gly Phe Thr Lys Thr Cys Gly Gly Pro Pro Cys Asn Ile  
 195 200 205  
 Gly Gly Ala Gly Asn Asn Thr Leu Thr Cys Pro Thr Asp Cys Phe Arg  
 210 215 220  
 Lys His Pro Glu Ala Thr Tyr Ala Arg Cys Gly Ser Gly Pro Trp Leu  
 225 230 235 240  
 Thr Pro Arg Cys Met Val His Tyr Pro Tyr Arg Leu Trp His Tyr Pro  
 245 250 255  
 Cys Thr Val Asn Phe Thr Ile Phe Lys Val Arg Met Tyr Val Gly Gly  
 260 265 270  
 Val Glu His Arg Phe Glu Ala Ala Cys Asn Trp Thr Arg Gly Glu Arg  
 275 280 285  
 Cys Asp Leu Glu Asp Arg Asp Arg Ser Glu Leu Ser Pro Leu Leu Leu  
 290 295 300  
 Ser Thr Thr Gly Asp Arg Gly Gln Thr Pro Ser Pro Pro Ser Leu  
 305 310 315

&lt;210&gt; 45

<211> 1395  
 <212> DNA  
 <213> Hepatitis C virus

<220>  
 <221> CDS  
 <222> 1..1392

<220>  
 <221> mat\_peptide  
 <222> 1..1389  
 <400> 45

atg gtg gcg ggg gcc cat tgg gga gtc ctg gcg ggc ctc gcc tac tat	48
Met Val Ala Gly Ala His Trp Gly Val Leu Ala Gly Leu Ala Tyr Tyr	
1 5 10 15	
tcc atg gtg ggg aac tgg gct aag gtt ttg gtt gtg atg cta ctc ttt	96
Ser Met Val Gly Asn Trp Ala Lys Val Leu Val Val Met Leu Leu Phe	
20 25 30	
gcc ggc gtc gac ggg cat acc cgc gtg tca gga ggg gca gca gcc tcc	144
Ala Gly Val Asp Gly His Thr Arg Val Ser Gly Gly Ala Ala Ala Ser	
35 40 45	
gat acc agg ggc ctt gtg tcc ctc ttt agc ccc ggg tcg gct cag aaa	192
Asp Thr Arg Gly Leu Val Ser Leu Phe Ser Pro Gly Ser Ala Gln Lys	
50 55 60	
atc cag ctc gta aac acc aac ggc agt tgg cac atc aac agg act gcc	240
Ile Gln Leu Val Asn Thr Asn Gly Ser Trp His Ile Asn Arg Thr Ala	
65 70 75 80	
ctg aac tgc aac gac tcc ctc caa aca ggg ttc ttt gcc gca cta ttc	288
Leu Asn Cys Asn Asp Ser Leu Gln Thr Gly Phe Phe Ala Ala Leu Phe	
85 90 95	
tac aaa cac aaa ttc aac tcg tct gga tgc cca gag cgc ttg gcc agc	336
Tyr Lys His Lys Phe Asn Ser Ser Gly Cys Pro Glu Arg Leu Ala Ser	
100 105 110	
tgt cgc tcc atc gac aag ttc gct cag ggg tgg ggt ccc ctc act tac	384
Cys Arg Ser Ile Asp Lys Phe Ala Gln Gly Trp Gly Pro Leu Thr Tyr	
115 120 125	
act gag cct aac agc tcg gac cag agg ccc tac tgc tgg cac tac gcg	432
Thr Glu Pro Asn Ser Ser Asp Gln Arg Pro Tyr Cys Trp His Tyr Ala	
130 135 140	
cct cga ccg tgt ggt att gta ccc gcg tct cag gtg tgc ggt cca gtg	480
Pro Arg Pro Cys Gly Ile Val Pro Ala Ser Gln Val Cys Gly Pro Val	
145 150 155 160	
tat tgc ttc acc ccg agc cct gtt gtg gtg ggg acg acc gat cgg ttt	528
Tyr Cys Phe Thr Pro Ser Pro Val Val Val Gly Thr Thr Asp Arg Phe	
165 170 175	
ggg gtc ccc acg tat aac tgg ggg gcg aac gac tcg gat gtg ctg att	576
Gly Val Pro Thr Tyr Asn Trp Gly Ala Asn Asp Ser Asp Val Leu Ile	
180 185 190	
ctc aac aac acg cgg ccg ccg cga ggc aac tgg ttc ggc tgt aca tgg	624

Leu	Asn	Asn	Thr	Arg	Pro	Pro	Arg	Gly	Asn	Trp	Phe	Gly	Cys	Thr	Trp	
		195					200					205				
atg	aat	ggc	act	ggg	ttc	acc	aag	acg	tgt	ggg	ggc	ccc	ccg	tgc	aac	672
Met	Asn	Gly	Thr	Gly	Phe	Thr	Lys	Thr	Cys	Gly	Gly	Pro	Pro	Cys	Asn	
	210					215				220						
atc	ggg	ggg	gcc	ggc	aac	aac	acc	ttg	acc	tgc	ccc	act	gac	tgt	ttt	720
Ile	Gly	Gly	Ala	Gly	Asn	Asn	Thr	Leu	Thr	Cys	Pro	Thr	Asp	Cys	Phe	
225					230					235					240	
cgg	aag	cac	ccc	gag	gcc	acc	tac	gcc	aga	tgc	ggt	tct	ggg	ccc	tgg	768
Arg	Lys	His	Pro	Glu	Ala	Thr	Tyr	Ala	Arg	Cys	Gly	Ser	Gly	Pro	Trp	
			245					250						255		
ctg	aca	cct	agg	tgt	atg	gtt	cat	tac	cca	tat	agg	ctc	tgg	cac	tac	816
Leu	Thr	Pro	Arg	Cys	Met	Val	His	Tyr	Pro	Tyr	Arg	Leu	Trp	His	Tyr	
			260					265					270			
ccc	tgc	act	gtc	aac	ttc	acc	atc	ttc	aag	gtt	agg	atg	tac	gtg	ggg	864
Pro	Cys	Thr	Val	Asn	Phe	Thr	Ile	Phe	Lys	Val	Arg	Met	Tyr	Val	Gly	
		275					280					285				
ggc	gtg	gag	cac	agg	ttc	gaa	gcc	gca	tgc	aat	tgg	act	cga	gga	gag	912
Gly	Val	Glu	His	Arg	Phe	Glu	Ala	Ala	Cys	Asn	Trp	Thr	Arg	Gly	Glu	
	290					295					300					
cgt	tgt	gac	ttg	gag	gac	agg	gat	aga	tca	gag	ctt	agc	ccg	ctg	ctg	960
Arg	Cys	Asp	Leu	Glu	Asp	Arg	Asp	Arg	Ser	Glu	Leu	Ser	Pro	Leu	Leu	
305					310					315					320	
ctg	tct	aca	aca	gag	tgg	cag	ata	ctg	ccc	tgt	tcc	ttc	acc	acc	ctg	1008
Leu	Ser	Thr	Thr	Glu	Trp	Gln	Ile	Leu	Pro	Cys	Ser	Phe	Thr	Thr	Leu	
				325					330					335		
ccg	gcc	cta	tcc	acc	ggc	ctg	atc	cac	ctc	cat	cag	aac	atc	gtg	gac	1056
Pro	Ala	Leu	Ser	Thr	Gly	Leu	Ile	His	Leu	His	Gln	Asn	Ile	Val	Asp	
			340					345					350			
gtg	caa	tac	ctg	tac	ggc	gta	ggg	tgc	gcg	gtt	gtc	tcc	ctt	gtc	atc	1104
Val	Gln	Tyr	Leu	Tyr	Gly	Val	Gly	Ser	Ala	Val	Val	Ser	Leu	Val	Ile	
		355				360						365				
aaa	tgg	gag	tat	gtc	ctg	ttg	ctc	ttc	ctt	ctc	ctg	gca	gac	gcg	cgc	1152
Lys	Trp	Glu	Tyr	Val	Leu	Leu	Leu	Phe	Leu	Leu	Leu	Ala	Asp	Ala	Arg	
	370					375					380					
atc	tgc	gcc	tgc	tta	tgg	atg	atg	ctg	ctg	ata	gct	caa	gct	gag	gcc	1200
Ile	Cys	Ala	Cys	Leu	Trp	Met	Met	Leu	Leu	Ile	Ala	Gln	Ala	Glu	Ala	
385					390					395					400	
gcc	tta	gag	aac	ctg	gtg	gtc	ctc	aat	gcg	gcg	gcc	gtg	gcc	ggg	gcg	1248
Ala	Leu	Glu	Asn	Leu	Val	Val	Leu	Asn	Ala	Ala	Ala	Val	Ala	Gly	Ala	
			405					410						415		
cat	ggc	act	ctt	tcc	ttc	ctt	gtg	ttc	ttc	tgt	gct	gcc	tgg	tac	atc	1296
His	Gly	Thr	Leu	Ser	Phe	Leu	Val	Phe	Phe	Cys	Ala	Ala	Trp	Tyr	Ile	
			420					425					430			
aag	ggc	agg	ctg	gtc	cct	ggt	gcg	gca	tac	gcc	ttc	tat	ggc	gtg	tgg	1344
Lys	Gly	Arg	Leu	Val	Pro	Gly	Ala	Ala	Tyr	Ala	Phe	Tyr	Gly	Val	Trp	

435

440

445

ccg ctg ctc ctg ctt ctg ctg gcc tta cca cca cga gct tat gcc tagtaa 1395  
 Pro Leu Leu Leu Leu Leu Leu Ala Leu Pro Pro Arg Ala Tyr Ala  
 450 455 460

&lt;210&gt; 46

&lt;211&gt; 463

&lt;212&gt; PRT

&lt;213&gt; Hepatitis C virus

&lt;400&gt; 46

Met Val Ala Gly Ala His Trp Gly Val Leu Ala Gly Leu Ala Tyr Tyr  
 1 5 10 15

Ser Met Val Gly Asn Trp Ala Lys Val Leu Val Val Met Leu Leu Phe  
 20 25 30

Ala Gly Val Asp Gly His Thr Arg Val Ser Gly Gly Ala Ala Ala Ser  
 35 40 45

Asp Thr Arg Gly Leu Val Ser Leu Phe Ser Pro Gly Ser Ala Gln Lys  
 50 55 60

Ile Gln Leu Val Asn Thr Asn Gly Ser Trp His Ile Asn Arg Thr Ala  
 65 70 75 80

Leu Asn Cys Asn Asp Ser Leu Gln Thr Gly Phe Phe Ala Ala Leu Phe  
 85 90 95

Tyr Lys His Lys Phe Asn Ser Ser Gly Cys Pro Glu Arg Leu Ala Ser  
 100 105 110

Cys Arg Ser Ile Asp Lys Phe Ala Gln Gly Trp Gly Pro Leu Thr Tyr  
 115 120 125

Thr Glu Pro Asn Ser Ser Asp Gln Arg Pro Tyr Cys Trp His Tyr Ala  
 130 135 140

Pro Arg Pro Cys Gly Ile Val Pro Ala Ser Gln Val Cys Gly Pro Val  
 145 150 155 160

Tyr Cys Phe Thr Pro Ser Pro Val Val Val Gly Thr Thr Asp Arg Phe  
 165 170 175

Gly Val Pro Thr Tyr Asn Trp Gly Ala Asn Asp Ser Asp Val Leu Ile  
 180 185 190

Leu Asn Asn Thr Arg Pro Pro Arg Gly Asn Trp Phe Gly Cys Thr Trp  
 195 200 205

Met Asn Gly Thr Gly Phe Thr Lys Thr Cys Gly Gly Pro Pro Cys Asn  
 210 215 220

Ile Gly Gly Ala Gly Asn Asn Thr Leu Thr Cys Pro Thr Asp Cys Phe  
 225 230 235 240

Arg Lys His Pro Glu Ala Thr Tyr Ala Arg Cys Gly Ser Gly Pro Trp  
 245 250 255

Leu Thr Pro Arg Cys Met Val His Tyr Pro Tyr Arg Leu Trp His Tyr  
 260 265 270  
 Pro Cys Thr Val Asn Phe Thr Ile Phe Lys Val Arg Met Tyr Val Gly  
 275 280 285  
 Gly Val Glu His Arg Phe Glu Ala Ala Cys Asn Trp Thr Arg Gly Glu  
 290 295 300  
 Arg Cys Asp Leu Glu Asp Arg Asp Arg Ser Glu Leu Ser Pro Leu Leu  
 305 310 315 320  
 Leu Ser Thr Thr Glu Trp Gln Ile Leu Pro Cys Ser Phe Thr Thr Leu  
 325 330 335  
 Pro Ala Leu Ser Thr Gly Leu Ile His Leu His Gln Asn Ile Val Asp  
 340 345 350  
 Val Gln Tyr Leu Tyr Gly Val Gly Ser Ala Val Val Ser Leu Val Ile  
 355 360 365  
 Lys Trp Glu Tyr Val Leu Leu Leu Phe Leu Leu Leu Ala Asp Ala Arg  
 370 375 380  
 Ile Cys Ala Cys Leu Trp Met Met Leu Leu Ile Ala Gln Ala Glu Ala  
 385 390 395 400  
 Ala Leu Glu Asn Leu Val Val Leu Asn Ala Ala Ala Val Ala Gly Ala  
 405 410 415  
 His Gly Thr Leu Ser Phe Leu Val Phe Phe Cys Ala Ala Trp Tyr Ile  
 420 425 430  
 Lys Gly Arg Leu Val Pro Gly Ala Ala Tyr Ala Phe Tyr Gly Val Trp  
 435 440 445  
 Pro Leu Leu Leu Leu Leu Leu Ala Leu Pro Pro Arg Ala Tyr Ala  
 450 455 460

&lt;210&gt; 47

&lt;211&gt; 2082

&lt;212&gt; DNA

&lt;213&gt; Hepatitis C virus

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; 1..2079

&lt;220&gt;

&lt;221&gt; mat\_peptide

&lt;222&gt; 1..2076

&lt;400&gt; 47

aat ttg ggt aag gtc atc gat acc ctt aca tgc ggc ttc gcc gac ctc  
 Asn Leu Gly Lys Val Ile Asp Thr Leu Thr Cys Gly Phe Ala Asp Leu  
 1 5 10 15

48

gtg ggg tac att ccg ctc gtc ggc gcc ccc cta ggg ggc gct gcc agg  
 Val Gly Tyr Ile Pro Leu Val Gly Ala Pro Leu Gly Gly Ala Ala Arg  
 20 25 30

96



gcc ctg gcg cat ggc gtc cgg gtt ctg gag gac ggc gtg aac tat gca Ala Leu Ala His Gly Val Arg Val Leu Glu Asp Gly Val Asn Tyr Ala 35 40 45	144
aca ggg aat ttg ccc ggt tgc tct ttc tct atc ttc ctc ttg gct ttg Thr Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile Phe Leu Leu Ala Leu 50 55 60	192
ctg tcc tgt ctg acc gtt cca gct tcc gct tat gaa gtg cgc aac gtg Leu Ser Cys Leu Thr Val Pro Ala Ser Ala Tyr Glu Val Arg Asn Val 65 70 75 80	240
tcc ggg atg tac cat gtc acg aac gac tgc tcc aac tca agc att gtg Ser Gly Met Tyr His Val Thr Asn Asp Cys Ser Asn Ser Ser Ile Val 85 90 95	288
tat gag gca gcg gac atg atc atg cac acc ccc ggg tgc gtg ccc tgc Tyr Glu Ala Ala Asp Met Ile Met His Thr Pro Gly Cys Val Pro Cys 100 105 110	336
gtt cgg gag aac aac tct tcc cgc tgc tgg gta gcg ctc acc ccc acg Val Arg Glu Asn Asn Ser Ser Arg Cys Trp Val Ala Leu Thr Pro Thr 115 120 125	384
ctc gca gct agg aac gcc agc gtc ccc acc acg aca ata cga cgc cac Leu Ala Ala Arg Asn Ala Ser Val Pro Thr Thr Ile Arg Arg His 130 135 140	432
gtc gat ttg ctc gtt ggg gcg gct gct ttc tgt tcc gct atg tac gtg Val Asp Leu Leu Val Gly Ala Ala Ala Phe Cys Ser Ala Met Tyr Val 145 150 155 160	480
ggg gac ctc tgc gga tct gtc ttc ctc gtc tcc cag ctg ttc acc atc Gly Asp Leu Cys Gly Ser Val Phe Leu Val Ser Gln Leu Phe Thr Ile 165 170 175	528
tcg cct cgc cgg cat gag acg gtg cag gac tgc aat tgc tca atc tat Ser Pro Arg Arg His Glu Thr Val Gln Asp Cys Asn Cys Ser Ile Tyr 180 185 190	576
ccc ggc cac ata acg ggt cac cgt atg gct tgg gat atg atg atg aac Pro Gly His Ile Thr Gly His Arg Met Ala Trp Asp Met Met Met Asn 195 200 205	624
tgg tcg cct aca acg gcc ctg gtg gta tcg cag ctg ctc cgg atc cca Trp Ser Pro Thr Thr Ala Leu Val Val Ser Gln Leu Leu Arg Ile Pro 210 215 220	672
caa gct gtc gtg gac atg gtg gcg ggg gcc cat tgg gga gtc ctg gcg Gln Ala Val Val Asp Met Val Ala Gly Ala His Trp Gly Val Leu Ala 225 230 235 240	720
ggc ctc gcc tac tat tcc atg gtg ggg aac tgg gct aag gtt ttg gtt Gly Leu Ala Tyr Tyr Ser Met Val Gly Asn Trp Ala Lys Val Leu Val 245 250 255	768
gtg atg cta ctc ttt gcc ggc gtc gac ggg cat acc cgc gtg tca gga Val Met Leu Leu Phe Ala Gly Val Asp Gly His Thr Arg Val Ser Gly 260 265 270	816
ggg gca gca gcc tcc gat acc agg ggc ctt gtg tcc ctc ttt agc ccc	864

Gly	Ala	Ala	Ala	Ser	Asp	Thr	Arg	Gly	Leu	Val	Ser	Leu	Phe	Ser	Pro		
	275						280					285					
ggg	tcg	gct	cag	aaa	atc	cag	ctc	gta	aac	acc	aac	ggc	agt	tgg	cac		912
Gly	Ser	Ala	Gln	Lys	Ile	Gln	Leu	Val	Asn	Thr	Asn	Gly	Ser	Trp	His		
	290					295					300						
atc	aac	agg	act	gcc	ctg	aac	tgc	aac	gac	tcc	ctc	caa	aca	ggg	ttc		960
Ile	Asn	Arg	Thr	Ala	Leu	Asn	Cys	Asn	Asp	Ser	Leu	Gln	Thr	Gly	Phe		
305					310					315					320		
ttt	gcc	gca	cta	ttc	tac	aaa	cac	aaa	ttc	aac	tcg	tct	gga	tgc	cca		1008
Phe	Ala	Ala	Leu	Phe	Tyr	Lys	His	Lys	Phe	Asn	Ser	Ser	Gly	Cys	Pro		
				325					330					335			
gag	cgc	ttg	gcc	agc	tgt	cgc	tcc	atc	gac	aag	ttc	gct	cag	ggg	tgg		1056
Glu	Arg	Leu	Ala	Ser	Cys	Arg	Ser	Ile	Asp	Lys	Phe	Ala	Gln	Gly	Trp		
			340					345					350				
ggt	ccc	ctc	act	tac	act	gag	cct	aac	agc	tcg	gac	cag	agg	ccc	tac		1104
Gly	Pro	Leu	Thr	Tyr	Thr	Glu	Pro	Asn	Ser	Ser	Asp	Gln	Arg	Pro	Tyr		
		355					360					365					
tgc	tgg	cac	tac	gcg	cct	cga	ccg	tgt	ggt	att	gta	ccc	gcg	tct	cag		1152
Cys	Trp	His	Tyr	Ala	Pro	Arg	Pro	Cys	Gly	Ile	Val	Pro	Ala	Ser	Gln		
	370					375					380						
gtg	tgc	ggt	cca	gtg	tat	tgc	ttc	acc	ccg	agc	cct	gtt	gtg	gtg	ggg		1200
Val	Cys	Gly	Pro	Val	Tyr	Cys	Phe	Thr	Pro	Ser	Pro	Val	Val	Val	Gly		
385					390					395					400		
acg	acc	gat	cgg	ttt	ggt	gtc	ccc	acg	tat	aac	tgg	ggg	gcg	aac	gac		1248
Thr	Thr	Asp	Arg	Phe	Gly	Val	Pro	Thr	Tyr	Asn	Trp	Gly	Ala	Asn	Asp		
				405					410					415			
tcg	gat	gtg	ctg	att	ctc	aac	aac	acg	cgg	ccg	ccg	cga	ggc	aac	tgg		1296
Ser	Asp	Val	Leu	Ile	Leu	Asn	Asn	Thr	Arg	Pro	Pro	Arg	Gly	Asn	Trp		
			420					425					430				
ttc	ggc	tgt	aca	tgg	atg	aat	ggc	act	ggg	ttc	acc	aag	acg	tgt	ggg		1344
Phe	Gly	Cys	Thr	Trp	Met	Asn	Gly	Thr	Gly	Phe	Thr	Lys	Thr	Cys	Gly		
			435				440					445					
ggc	ccc	ccg	tgc	aac	atc	ggg	ggg	gcc	ggc	aac	aac	acc	ttg	acc	tgc		1392
Gly	Pro	Pro	Cys	Asn	Ile	Gly	Gly	Ala	Gly	Asn	Asn	Thr	Leu	Thr	Cys		
	450					455					460						
ccc	act	gac	tgt	ttt	cgg	aag	cac	ccc	gag	gcc	acc	tac	gcc	aga	tgc		1440
Pro	Thr	Asp	Cys	Phe	Arg	Lys	His	Pro	Glu	Ala	Thr	Tyr	Ala	Arg	Cys		
					470					475				480			
ggt	tct	ggg	ccc	tgg	ctg	aca	cct	agg	tgt	atg	gtt	cat	tac	cca	tat		1488
Gly	Ser	Gly	Pro	Trp	Leu	Thr	Pro	Arg	Cys	Met	Val	His	Tyr	Pro	Tyr		
				485					490					495			
agg	ctc	tgg	cac	tac	ccc	tgc	act	gtc	aac	ttc	acc	atc	ttc	aag	gtt		1536
Arg	Leu	Trp	His	Tyr	Pro	Cys	Thr	Val	Asn	Phe	Thr	Ile	Phe	Lys	Val		
			500					505				510					
agg	atg	tac	gtg	ggg	ggc	gtg	gag	cac	agg	ttc	gaa	gcc	gca	tgc	aat		1584
Arg	Met	Tyr	Val	Gly	Gly	Val	Glu	His	Arg	Phe	Glu	Ala	Ala	Cys	Asn		
		515					520					525					

tgg act cga gga gag cgt tgt gac ttg gag gac agg gat aga tca gag 1632  
 Trp Thr Arg Gly Glu Arg Cys Asp Leu Glu Asp Arg Asp Arg Ser Glu  
 530 535 540

ctt agc ccg ctg ctg ctg tct aca aca gag ttg cag ata ctg ccc tgt 1680  
 Leu Ser Pro Leu Leu Leu Ser Thr Thr Glu Trp Gln Ile Leu Pro Cys  
 545 550 555 560

tcc ttc acc acc ctg ccg gcc cta tcc acc ggc ctg atc cac ctc cat 1728  
 Ser Phe Thr Thr Leu Pro Ala Leu Ser Thr Gly Leu Ile His Leu His  
 565 570 575

cag aac atc gtg gac gtg caa tac ctg tac ggt gta ggg tcg gcg gtt 1776  
 Gln Asn Ile Val Asp Val Gln Tyr Leu Tyr Gly Val Gly Ser Ala Val  
 580 585 590

gtc tcc ctt gtc atc aaa ttg gag tat gtc ctg ttg ctc ttc ctt ctc 1824  
 Val Ser Leu Val Ile Lys Trp Glu Tyr Val Leu Leu Leu Phe Leu Leu  
 595 600 605

ctg gca gac gcg cgc atc tgc gcc tgc tta ttg atg atg ctg ctg ata 1872  
 Leu Ala Asp Ala Arg Ile Cys Ala Cys Leu Trp Met Met Leu Leu Ile  
 610 615 620

gct caa gct gag gcc gcc tta gag aac ctg gtg gtc ctc aat gcg gcg 1920  
 Ala Gln Ala Glu Ala Ala Leu Glu Asn Leu Val Val Leu Asn Ala Ala  
 625 630 635 640

gcc gtg gcc ggg gcg cat ggc act ctt tcc ttc ctt gtg ttc ttc tgt 1968  
 Ala Val Ala Gly Ala His Gly Thr Leu Ser Phe Leu Val Phe Phe Cys  
 645 650 655

gct gcc tgg tac atc aag ggc agg ctg gtc cct ggt gcg gca tac gcc 2016  
 Ala Ala Trp Tyr Ile Lys Gly Arg Leu Val Pro Gly Ala Ala Tyr Ala  
 660 665 670

ttc tat ggc gtg tgg ccg ctg ctc ctg ctt ctg ctg gcc tta cca cca 2064  
 Phe Tyr Gly Val Trp Pro Leu Leu Leu Leu Leu Ala Leu Pro Pro  
 675 680 685

cga gct tat gcc tagtaa 2082  
 Arg Ala Tyr Ala  
 690

<210> 48

<211> 692  
 <212> PRT  
 <213> Hepatitis C virus

<400> 48

Asn Leu Gly Lys Val Ile Asp Thr Leu Thr Cys Gly Phe Ala Asp Leu  
 1 5 10 15

Val Gly Tyr Ile Pro Leu Val Gly Ala Pro Leu Gly Gly Ala Ala Arg  
 20 25 30

Ala Leu Ala His Gly Val Arg Val Leu Glu Asp Gly Val Asn Tyr Ala  
 35 40 45

Thr Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile Phe Leu Leu Ala Leu  
 50 55 60  
 Leu Ser Cys Leu Thr Val Pro Ala Ser Ala Tyr Glu Val Arg Asn Val  
 65 70 75 80  
 Ser Gly Met Tyr His Val Thr Asn Asp Cys Ser Asn Ser Ser Ile Val  
 85 90 95  
 Tyr Glu Ala Ala Asp Met Ile Met His Thr Pro Gly Cys Val Pro Cys  
 100 105 110  
 Val Arg Glu Asn Asn Ser Ser Arg Cys Trp Val Ala Leu Thr Pro Thr  
 115 120 125  
 Leu Ala Ala Arg Asn Ala Ser Val Pro Thr Thr Thr Ile Arg Arg His  
 130 135 140  
 Val Asp Leu Leu Val Gly Ala Ala Ala Phe Cys Ser Ala Met Tyr Val  
 145 150 155 160  
 Gly Asp Leu Cys Gly Ser Val Phe Leu Val Ser Gln Leu Phe Thr Ile  
 165 170 175  
 Ser Pro Arg Arg His Glu Thr Val Gln Asp Cys Asn Cys Ser Ile Tyr  
 180 185 190  
 Pro Gly His Ile Thr Gly His Arg Met Ala Trp Asp Met Met Met Asn  
 195 200 205  
 Trp Ser Pro Thr Thr Ala Leu Val Val Ser Gln Leu Leu Arg Ile Pro  
 210 215 220  
 Gln Ala Val Val Asp Met Val Ala Gly Ala His Trp Gly Val Leu Ala  
 225 230 235 240  
 Gly Leu Ala Tyr Tyr Ser Met Val Gly Asn Trp Ala Lys Val Leu Val  
 245 250 255  
 Val Met Leu Leu Phe Ala Gly Val Asp Gly His Thr Arg Val Ser Gly  
 260 265 270  
 Gly Ala Ala Ala Ser Asp Thr Arg Gly Leu Val Ser Leu Phe Ser Pro  
 275 280 285  
 Gly Ser Ala Gln Lys Ile Gln Leu Val Asn Thr Asn Gly Ser Trp His  
 290 295 300  
 Ile Asn Arg Thr Ala Leu Asn Cys Asn Asp Ser Leu Gln Thr Gly Phe  
 305 310 315 320  
 Phe Ala Ala Leu Phe Tyr Lys His Lys Phe Asn Ser Ser Gly Cys Pro  
 325 330 335  
 Glu Arg Leu Ala Ser Cys Arg Ser Ile Asp Lys Phe Ala Gln Gly Trp  
 340 345 350  
 Gly Pro Leu Thr Tyr Thr Glu Pro Asn Ser Ser Asp Gln Arg Pro Tyr  
 355 360 365  
 Cys Trp His Tyr Ala Pro Arg Pro Cys Gly Ile Val Pro Ala Ser Gln

370 375 380  
 Val Cys Gly Pro Val Tyr Cys Phe Thr Pro Ser Pro Val Val Val Gly  
 385 390 395 400  
 Thr Thr Asp Arg Phe Gly Val Pro Thr Tyr Asn Trp Gly Ala Asn Asp  
 405 410 415  
 Ser Asp Val Leu Ile Leu Asn Asn Thr Arg Pro Pro Arg Gly Asn Trp  
 420 425 430  
 Phe Gly Cys Thr Trp Met Asn Gly Thr Gly Phe Thr Lys Thr Cys Gly  
 435 440 445  
 Gly Pro Pro Cys Asn Ile Gly Gly Ala Gly Asn Asn Thr Leu Thr Cys  
 450 455 460  
 Pro Thr Asp Cys Phe Arg Lys His Pro Glu Ala Thr Tyr Ala Arg Cys  
 465 470 475 480  
 Gly Ser Gly Pro Trp Leu Thr Pro Arg Cys Met Val His Tyr Pro Tyr  
 485 490 495  
 Arg Leu Trp His Tyr Pro Cys Thr Val Asn Phe Thr Ile Phe Lys Val  
 500 505 510  
 Arg Met Tyr Val Gly Gly Val Glu His Arg Phe Glu Ala Ala Cys Asn  
 515 520 525  
 Trp Thr Arg Gly Glu Arg Cys Asp Leu Glu Asp Arg Asp Arg Ser Glu  
 530 535 540  
 Leu Ser Pro Leu Leu Leu Ser Thr Thr Glu Trp Gln Ile Leu Pro Cys  
 545 550 555 560  
 Ser Phe Thr Thr Leu Pro Ala Leu Ser Thr Gly Leu Ile His Leu His  
 565 570 575  
 Gln Asn Ile Val Asp Val Gln Tyr Leu Tyr Gly Val Gly Ser Ala Val  
 580 585 590  
 Val Ser Leu Val Ile Lys Trp Glu Tyr Val Leu Leu Leu Phe Leu Leu  
 595 600 605  
 Leu Ala Asp Ala Arg Ile Cys Ala Cys Leu Trp Met Met Leu Leu Ile  
 610 615 620  
 Ala Gln Ala Glu Ala Ala Leu Glu Asn Leu Val Val Leu Asn Ala Ala  
 625 630 635 640  
 Ala Val Ala Gly Ala His Gly Thr Leu Ser Phe Leu Val Phe Phe Cys  
 645 650 655  
 Ala Ala Trp Tyr Ile Lys Gly Arg Leu Val Pro Gly Ala Ala Tyr Ala  
 660 665 670  
 Phe Tyr Gly Val Trp Pro Leu Leu Leu Leu Leu Ala Leu Pro Pro  
 675 680 685  
 Arg Ala Tyr Ala  
 690

&lt;210&gt; 49

&lt;211&gt; 2433

&lt;212&gt; DNA

&lt;213&gt; Hepatitis C virus

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; 1..2430

&lt;220&gt;

&lt;221&gt; mat\_peptide

&lt;222&gt; 1..2427

&lt;400&gt; 49

atg agc acg aat cct aaa cct caa aga aaa acc aaa cgt aac acc aac	48
Met Ser Thr Asn Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn	
1 5 10 15	
cgc cgc cca cag gac gtc aag ttc ccg ggc ggt ggt cag atc gtt ggt	96
Arg Arg Pro Gln Asp Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly	
20 25 30	
gga gtt tac ctg ttg ccg cgc agg ggc ccc agg ttg ggt gtg cgc gcg	144
Gly Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg Leu Gly Val Arg Ala	
35 40 45	
act agg aag act tcc gag cgg tcg caa cct cgt ggg agg cga caa cct	192
Thr Arg Lys Thr Ser Glu Arg Ser Gln Pro Arg Gly Arg Arg Gln Pro	
50 55 60	
atc ccc aag gct cgc cga ccc gag ggt agg gcc tgg gct cag ccc ggg	240
Ile Pro Lys Ala Arg Arg Pro Glu Gly Arg Ala Trp Ala Gln Pro Gly	
65 70 75 80	
tac cct tgg ccc ctc tat ggc aat gag ggc atg ggg tgg gca gga tgg	288
Tyr Pro Trp Pro Leu Tyr Gly Asn Glu Gly Met Gly Trp Ala Gly Trp	
85 90 95	
ctc ctg tca ccc cgc ggc tct cgg cct agt tgg ggc cct aca gac ccc	336
Leu Leu Ser Pro Arg Gly Ser Arg Pro Ser Trp Gly Pro Thr Asp Pro	
100 105 110	
cgg cgt agg tcg cgt aat ttg ggt aag gtc atc gat acc ctt aca tgc	384
Arg Arg Arg Ser Arg Asn Leu Gly Lys Val Ile Asp Thr Leu Thr Cys	
115 120 125	
ggc ttc gcc gac ctc gtg ggg tac att ccg ctc gtc ggc gcc ccc cta	432
Gly Phe Ala Asp Leu Val Gly Tyr Ile Pro Leu Val Gly Ala Pro Leu	
130 135 140	
ggg ggc gct gcc agg gcc ctg gcg cat ggc gtc cgg gtt ctg gag gac	480
Gly Gly Ala Ala Arg Ala Leu Ala His Gly Val Arg Val Leu Glu Asp	
145 150 155 160	
ggc gtg aac tat gca aca ggg aat ttg ccc ggt tgc tct ttc tct atc	528
Gly Val Asn Tyr Ala Thr Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile	
165 170 175	
ttc ctc ttg gct ttg ctg tcc tgt ctg acc gtt cca gct tcc gct tat	576

Phe	Leu	Leu	Ala	Leu	Leu	Ser	Cys	Leu	Thr	Val	Pro	Ala	Ser	Ala	Tyr	
			180					185					190			
gaa	gtg	cgc	aac	gtg	tcc	ggg	atg	tac	cat	gtc	acg	aac	gac	tgc	tcc	624
Glu	Val	Arg	Asn	Val	Ser	Gly	Met	Tyr	His	Val	Thr	Asn	Asp	Cys	Ser	
			195				200					205				
aac	tca	agc	att	gtg	tat	gag	gca	gcg	gac	atg	atc	atg	cac	acc	ccc	672
Asn	Ser	Ser	Ile	Val	Tyr	Glu	Ala	Ala	Asp	Met	Ile	Met	His	Thr	Pro	
			210			215				220						
ggg	tgc	gtg	ccc	tgc	gtt	cgg	gag	aac	aac	tct	tcc	cgc	tgc	tgg	gta	720
Gly	Cys	Val	Pro	Cys	Val	Arg	Glu	Asn	Asn	Ser	Ser	Arg	Cys	Trp	Val	
225					230					235					240	
gcg	ctc	acc	ccc	acg	ctc	gca	gct	agg	aac	gcc	agc	gtc	ccc	acc	acg	768
Ala	Leu	Thr	Pro	Thr	Leu	Ala	Ala	Arg	Asn	Ala	Ser	Val	Pro	Thr	Thr	
				245					250					255		
aca	ata	cga	cgc	cac	gtc	gat	ttg	ctc	gtt	ggg	gcg	gct	gct	ttc	tgt	816
Thr	Ile	Arg	Arg	His	Val	Asp	Leu	Leu	Val	Gly	Ala	Ala	Ala	Phe	Cys	
				260				265					270			
tcc	gct	atg	tac	gtg	ggg	gac	ctc	tgc	gga	tct	gtc	ttc	ctc	gtc	tcc	864
Ser	Ala	Met	Tyr	Val	Gly	Asp	Leu	Cys	Gly	Ser	Val	Phe	Leu	Val	Ser	
			275				280					285				
cag	ctg	ttc	acc	atc	tcg	cct	cgc	cgg	cat	gag	acg	gtg	cag	gac	tgc	912
Gln	Leu	Phe	Thr	Ile	Ser	Pro	Arg	Arg	His	Glu	Thr	Val	Gln	Asp	Cys	
			290			295					300					
aat	tgc	tca	atc	tat	ccc	ggc	cac	ata	acg	ggt	cac	cgt	atg	gct	tgg	960
Asn	Cys	Ser	Ile	Tyr	Pro	Gly	His	Ile	Thr	Gly	His	Arg	Met	Ala	Trp	
305					310					315				320		
gat	atg	atg	atg	aac	tgg	tcg	cct	aca	acg	gcc	ctg	gtg	gta	tcg	cag	1008
Asp	Met	Met	Met	Asn	Trp	Ser	Pro	Thr	Thr	Ala	Leu	Val	Val	Ser	Gln	
				325					330					335		
ctg	ctc	cgg	atc	cca	caa	gct	gtc	gtg	gac	atg	gtg	gcg	ggg	gcc	cat	1056
Leu	Leu	Arg	Ile	Pro	Gln	Ala	Val	Val	Asp	Met	Val	Ala	Gly	Ala	His	
			340				345					350				
tgg	gga	gtc	ctg	gcg	ggc	ctc	gcc	tac	tat	tcc	atg	gtg	ggg	aac	tgg	1104
Trp	Gly	Val	Leu	Ala	Gly	Leu	Ala	Tyr	Tyr	Ser	Met	Val	Gly	Asn	Trp	
			355			360						365				
gct	aag	gtt	ttg	gtt	gtg	atg	cta	ctc	ttt	gcc	ggc	gtc	gac	ggg	cat	1152
Ala	Lys	Val	Leu	Val	Val	Met	Leu	Leu	Phe	Ala	Gly	Val	Asp	Gly	His	
			370			375					380					
acc	cgc	gtg	tca	gga	ggg	gca	gca	gcc	tcc	gat	acc	agg	ggc	ctt	gtg	1200
Thr	Arg	Val	Ser	Gly	Gly	Ala	Ala	Ala	Ser	Asp	Thr	Arg	Gly	Leu	Val	
385					390					395					400	
tcc	ctc	ttt	agc	ccc	ggg	tcg	gct	cag	aaa	atc	cag	ctc	gta	aac	acc	1248
Ser	Leu	Phe	Ser	Pro	Gly	Ser	Ala	Gln	Lys	Ile	Gln	Leu	Val	Asn	Thr	
				405					410					415		
aac	ggc	agt	tgg	cac	atc	aac	agg	act	gcc	ctg	aac	tgc	aac	gac	tcc	1296
Asn	Gly	Ser	Trp	His	Ile	Asn	Arg	Thr	Ala	Leu	Asn	Cys	Asn	Asp	Ser	

420	425	430	
ctc caa aca ggg ttc ttt gcc gca cta ttc tac aaa cac aaa ttc aac Leu Gln Thr Gly Phe Phe Ala Ala Leu Phe Tyr Lys His Lys Phe Asn 435 440 445			1344
tcg tct gga tgc cca gag cgc ttg gcc agc tgt cgc tcc atc gac aag Ser Ser Gly Cys Pro Glu Arg Leu Ala Ser Cys Arg Ser Ile Asp Lys 450 455 460			1392
ttc gct cag ggg tgg ggt ccc ctc act tac act gag cct aac agc tcg Phe Ala Gln Gly Trp Gly Pro Leu Thr Tyr Thr Glu Pro Asn Ser Ser 465 470 475 480			1440
gac cag agg ccc tac tgc tgg cac tac gcg cct cga ccg tgt ggt att Asp Gln Arg Pro Tyr Cys Trp His Tyr Ala Pro Arg Pro Cys Gly Ile 485 490 495			1488
gta ccc gcg tct cag gtg tgc ggt cca gtg tat tgc ttc acc ccg agc Val Pro Ala Ser Gln Val Cys Gly Pro Val Tyr Cys Phe Thr Pro Ser 500 505 510			1536
cct gtt gtg gtg ggg acg acc gat cgg ttt ggt gtc ccc acg tat aac Pro Val Val Val Gly Thr Thr Asp Arg Phe Gly Val Pro Thr Tyr Asn 515 520 525			1584
tgg ggg gcg aac gac tcg gat gtg ctg att ctc aac aac acg cgg ccg Trp Gly Ala Asn Asp Ser Asp Val Leu Ile Leu Asn Asn Thr Arg Pro 530 535 540			1632
ccg cga ggc aac tgg ttc ggc tgt aca tgg atg aat ggc act ggg ttc Pro Arg Gly Asn Trp Phe Gly Cys Thr Trp Met Asn Gly Thr Gly Phe 545 550 555 560			1680
acc aag acg tgt ggg ggc ccc ccg tgc aac atc ggg ggg gcc ggc aac Thr Lys Thr Cys Gly Gly Pro Pro Cys Asn Ile Gly Gly Ala Gly Asn 565 570 575			1728
aac acc ttg acc tgc ccc act gac tgt ttt cgg aag cac ccc gag gcc Asn Thr Leu Thr Cys Pro Thr Asp Cys Phe Arg Lys His Pro Glu Ala 580 585 590			1776
acc tac gcc aga tgc ggt tct ggg ccc tgg ctg aca cct agg tgt atg Thr Tyr Ala Arg Cys Gly Ser Gly Pro Trp Leu Thr Pro Arg Cys Met 595 600 605			1824
gtt cat tac cca tat agg ctc tgg cac tac ccc tgc act gtc aac ttc Val His Tyr Pro Tyr Arg Leu Trp His Tyr Pro Cys Thr Val Asn Phe 610 615 620			1872
acc atc ttc aag gtt agg atg tac gtg ggg ggc gtg gag cac agg ttc Thr Ile Phe Lys Val Arg Met Tyr Val Gly Gly Val Glu His Arg Phe 625 630 635 640			1920
gaa gcc gca tgc aat tgg act cga gga gag cgt tgt gac ttg gag gac Glu Ala Ala Cys Asn Trp Thr Arg Gly Glu Arg Cys Asp Leu Glu Asp 645 650 655			1968
agg gat aga tca gag ctt agc ccg ctg ctg ctg tct aca aca gag tgg Arg Asp Arg Ser Glu Leu Ser Pro Leu Leu Leu Ser Thr Thr Glu Trp 660 665 670			2016



cag ata ctg ccc tgt tcc ttc acc acc ctg ccg gcc cta tcc acc ggc 2064  
 Gln Ile Leu Pro Cys Ser Phe Thr Thr Leu Pro Ala Leu Ser Thr Gly  
 675 680 685

ctg atc cac ctc cat cag aac atc gtg gac gtg caa tac ctg tac ggt 2112  
 Leu Ile His Leu His Gln Asn Ile Val Asp Val Gln Tyr Leu Tyr Gly  
 690 695 700

gta ggg tcg gcg gtt gtc tcc ctt gtc atc aaa tgg gag tat gtc ctg 2160  
 Val Gly Ser Ala Val Val Ser Leu Val Ile Lys Trp Glu Tyr Val Leu  
 705 710 715 720

ttg ctc ttc ctt ctc ctg gca gac gcg cgc atc tgc gcc tgc tta tgg 2208  
 Leu Leu Phe Leu Leu Leu Ala Asp Ala Arg Ile Cys Ala Cys Leu Trp  
 725 730 735

atg atg ctg ctg ata gct caa gct gag gcc gcc tta gag aac ctg gtg 2256  
 Met Met Leu Leu Ile Ala Gln Ala Glu Ala Ala Leu Glu Asn Leu Val  
 740 745 750

gtc ctc aat gcg gcg gcc gtg gcc ggg gcg cat ggc act ctt tcc ttc 2304  
 Val Leu Asn Ala Ala Ala Val Ala Gly Ala His Gly Thr Leu Ser Phe  
 755 760 765

ctt gtg ttc ttc tgt gct gcc tgg tac atc aag ggc agg ctg gtc cct 2352  
 Leu Val Phe Phe Cys Ala Ala Trp Tyr Ile Lys Gly Arg Leu Val Pro  
 770 775 780

ggg ggc gca tac gcc ttc tat ggc gtg tgg ccg ctg ctc ctg ctt ctg 2400  
 Gly Ala Ala Tyr Ala Phe Tyr Gly Val Trp Pro Leu Leu Leu Leu Leu  
 785 790 795 800

ctg gcc tta cca cca cga gct tat gcc tagtaa 2433  
 Leu Ala Leu Pro Pro Arg Ala Tyr Ala  
 805 810

&lt;210&gt; 50

&lt;211&gt; 809

&lt;212&gt; PRT

&lt;213&gt; Hepatitis C virus

&lt;400&gt; 50

Met Ser Thr Asn Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn  
 1 5 10 15

Arg Arg Pro Gln Asp Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly  
 20 25 30

Gly Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg Leu Gly Val Arg Ala  
 35 40 45

Thr Arg Lys Thr Ser Glu Arg Ser Gln Pro Arg Gly Arg Arg Gln Pro  
 50 55 60

Ile Pro Lys Ala Arg Arg Pro Glu Gly Arg Ala Trp Ala Gln Pro Gly  
 65 70 75 80

Tyr Pro Trp Pro Leu Tyr Gly Asn Glu Gly Met Gly Trp Ala Gly Trp

				85				90				95			
Leu	Leu	Ser	Pro	Arg	Gly	Ser	Arg	Pro	Ser	Trp	Gly	Pro	Thr	Asp	Pro
			100					105					110		
Arg	Arg	Arg	Ser	Arg	Asn	Leu	Gly	Lys	Val	Ile	Asp	Thr	Leu	Thr	Cys
		115					120					125			
Gly	Phe	Ala	Asp	Leu	Val	Gly	Tyr	Ile	Pro	Leu	Val	Gly	Ala	Pro	Leu
	130					135					140				
Gly	Gly	Ala	Ala	Arg	Ala	Leu	Ala	His	Gly	Val	Arg	Val	Leu	Glu	Asp
145					150					155				160	
Gly	Val	Asn	Tyr	Ala	Thr	Gly	Asn	Leu	Pro	Gly	Cys	Ser	Phe	Ser	Ile
				165					170					175	
Phe	Leu	Leu	Ala	Leu	Leu	Ser	Cys	Leu	Thr	Val	Pro	Ala	Ser	Ala	Tyr
			180					185					190		
Glu	Val	Arg	Asn	Val	Ser	Gly	Met	Tyr	His	Val	Thr	Asn	Asp	Cys	Ser
		195					200					205			
Asn	Ser	Ser	Ile	Val	Tyr	Glu	Ala	Ala	Asp	Met	Ile	Met	His	Thr	Pro
	210					215					220				
Gly	Cys	Val	Pro	Cys	Val	Arg	Glu	Asn	Asn	Ser	Ser	Arg	Cys	Trp	Val
225					230					235					240
Ala	Leu	Thr	Pro	Thr	Leu	Ala	Ala	Arg	Asn	Ala	Ser	Val	Pro	Thr	Thr
				245					250					255	
Thr	Ile	Arg	Arg	His	Val	Asp	Leu	Leu	Val	Gly	Ala	Ala	Ala	Phe	Cys
			260					265					270		
Ser	Ala	Met	Tyr	Val	Gly	Asp	Leu	Cys	Gly	Ser	Val	Phe	Leu	Val	Ser
		275					280					285			
Gln	Leu	Phe	Thr	Ile	Ser	Pro	Arg	Arg	His	Glu	Thr	Val	Gln	Asp	Cys
	290					295					300				
Asn	Cys	Ser	Ile	Tyr	Pro	Gly	His	Ile	Thr	Gly	His	Arg	Met	Ala	Trp
305					310					315					320
Asp	Met	Met	Met	Asn	Trp	Ser	Pro	Thr	Thr	Ala	Leu	Val	Val	Ser	Gln
				325					330					335	
Leu	Leu	Arg	Ile	Pro	Gln	Ala	Val	Val	Asp	Met	Val	Ala	Gly	Ala	His
			340					345					350		
Trp	Gly	Val	Leu	Ala	Gly	Leu	Ala	Tyr	Tyr	Ser	Met	Val	Gly	Asn	Trp
		355					360					365			
Ala	Lys	Val	Leu	Val	Val	Met	Leu	Leu	Phe	Ala	Gly	Val	Asp	Gly	His
						375					380				
Thr	Arg	Val	Ser	Gly	Gly	Ala	Ala	Ala	Ser	Asp	Thr	Arg	Gly	Leu	Val
385					390					395					400
Ser	Leu	Phe	Ser	Pro	Gly	Ser	Ala	Gln	Lys	Ile	Gln	Leu	Val	Asn	Thr
				405					410					415	

Asn Gly Ser Trp His Ile Asn Arg Thr Ala Leu Asn Cys Asn Asp Ser  
 420 425 430  
 Leu Gln Thr Gly Phe Phe Ala Ala Leu Phe Tyr Lys His Lys Phe Asn  
 435 440 445  
 Ser Ser Gly Cys Pro Glu Arg Leu Ala Ser Cys Arg Ser Ile Asp Lys  
 450 455 460  
 Phe Ala Gln Gly Trp Gly Pro Leu Thr Tyr Thr Glu Pro Asn Ser Ser  
 465 470 475 480  
 Asp Gln Arg Pro Tyr Cys Trp His Tyr Ala Pro Arg Pro Cys Gly Ile  
 485 490 495  
 Val Pro Ala Ser Gln Val Cys Gly Pro Val Tyr Cys Phe Thr Pro Ser  
 500 505 510  
 Pro Val Val Val Gly Thr Thr Asp Arg Phe Gly Val Pro Thr Tyr Asn  
 515 520 525  
 Trp Gly Ala Asn Asp Ser Asp Val Leu Ile Leu Asn Asn Thr Arg Pro  
 530 535 540  
 Pro Arg Gly Asn Trp Phe Gly Cys Thr Trp Met Asn Gly Thr Gly Phe  
 545 550 555 560  
 Thr Lys Thr Cys Gly Gly Pro Pro Cys Asn Ile Gly Gly Ala Gly Asn  
 565 570 575  
 Asn Thr Leu Thr Cys Pro Thr Asp Cys Phe Arg Lys His Pro Glu Ala  
 580 585 590  
 Thr Tyr Ala Arg Cys Gly Ser Gly Pro Trp Leu Thr Pro Arg Cys Met  
 595 600 605  
 Val His Tyr Pro Tyr Arg Leu Trp His Tyr Pro Cys Thr Val Asn Phe  
 610 615 620  
 Thr Ile Phe Lys Val Arg Met Tyr Val Gly Gly Val Glu His Arg Phe  
 625 630 635 640  
 Glu Ala Ala Cys Asn Trp Thr Arg Gly Glu Arg Cys Asp Leu Glu Asp  
 645 650 655  
 Arg Asp Arg Ser Glu Leu Ser Pro Leu Leu Leu Ser Thr Thr Glu Trp  
 660 665 670  
 Gln Ile Leu Pro Cys Ser Phe Thr Thr Leu Pro Ala Leu Ser Thr Gly  
 675 680 685  
 Leu Ile His Leu His Gln Asn Ile Val Asp Val Gln Tyr Leu Tyr Gly  
 690 695 700  
 Val Gly Ser Ala Val Val Ser Leu Val Ile Lys Trp Glu Tyr Val Leu  
 705 710 715 720  
 Leu Leu Phe Leu Leu Leu Ala Asp Ala Arg Ile Cys Ala Cys Leu Trp  
 725 730 735  
 Met Met Leu Leu Ile Ala Gln Ala Glu Ala Ala Leu Glu Asn Leu Val

740

745

750

Val Leu Asn Ala Ala Ala Val Ala Gly Ala His Gly Thr Leu Ser Phe  
755 760 765

Leu Val Phe Phe Cys Ala Ala Trp Tyr Ile Lys Gly Arg Leu Val Pro  
770 775 780

Gly Ala Ala Tyr Ala Phe Tyr Gly Val Trp Pro Leu Leu Leu Leu Leu  
785 790 795 800

Leu Ala Leu Pro Pro Arg Ala Tyr Ala  
805

&lt;210&gt; 51

&lt;211&gt; 17

&lt;212&gt; PRT

&lt;213&gt; Hepatitis C virus

&lt;220&gt;

&lt;221&gt; Modified-site

&lt;222&gt; 1..17

&lt;400&gt; 51

Ser Asn Ser Ser Glu Ala Ala Asp Met Ile Met His Thr Pro Gly Cys  
1 5 10 15

Val

&lt;210&gt; 52

&lt;211&gt; 22

&lt;212&gt; PRT

&lt;213&gt; Hepatitis C virus

&lt;220&gt;

&lt;221&gt; Modified-site

&lt;222&gt; 1..22

&lt;400&gt; 52

Gly Gly Ile Thr Gly His Arg Met Ala Trp Asp Met Met Met Asn Trp  
1 5 10 15

Ser Pro Thr Thr Ala Leu  
20

&lt;210&gt; 53

&lt;211&gt; 37

&lt;212&gt; PRT

&lt;213&gt; Hepatitis C virus

&lt;220&gt;

&lt;221&gt; Modified-site

&lt;222&gt; 1..37

&lt;400&gt; 53

Tyr Glu Val Arg Asn Val Ser Gly Ile Tyr His Val Thr Asn Asp Cys  
1 5 10 15

Ser Asn Ser Ser Ile Val Tyr Glu Ala Ala Asp Met Ile Met His Thr  
20 25 30

Pro Gly Cys Gly Lys  
35

<210> 54

<211> 25

<212> PRT

<213> Hepatitis C virus

<220>

<221> Modified-site

<222> 1..25

<400> 54

Gly Gly Thr Pro Thr Val Ala Thr Arg Asp Gly Lys Leu Pro Ala Thr  
1 5 10 15

Gln Leu Arg Arg His Ile Asp Leu Leu  
20 25

<210> 55

<211> 25

<212> PRT

<213> Hepatitis C virus

<220>

<221> Modified-site

<222> 1..25

<400> 55

Gly Gly Thr Pro Thr Leu Ala Ala Arg Asp Ala Ser Val Pro Thr Thr  
1 5 10 15

Thr Ile Arg Arg His Val Asp Leu Leu  
20 25

<210> 56

<211> 20

<212> PRT

<213> Hepatitis C virus

<400> 56

Leu Leu Ser Cys Leu Thr Val Pro Ala Ser Ala Tyr Gln Val Arg Asn  
1 5 10 15

Ser Thr Gly Leu  
20

<210> 57

<211> 20

<212> PRT

<213> Hepatitis C virus

&lt;400&gt; 57

Gln Val Arg Asn Ser Thr Gly Leu Tyr His Val Thr Asn Asp Cys Pro  
1 5 10 15

Asn Ser Ser Ile  
20

&lt;210&gt; 58

&lt;211&gt; 20

&lt;212&gt; PRT

&lt;213&gt; Hepatitis C virus

&lt;400&gt; 58

Asn Asp Cys Pro Asn Ser Ser Ile Val Tyr Glu Ala His Asp Ala Ile  
1 5 10 15

Leu His Thr Pro  
20

&lt;210&gt; 59

&lt;211&gt; 20

&lt;212&gt; PRT

&lt;213&gt; Hepatitis C virus

&lt;400&gt; 59

Ser Asn Ser Ser Ile Val Tyr Glu Ala Ala Asp Met Ile Met His Thr  
1 5 10 15

Pro Gly Cys Val  
20

&lt;210&gt; 60

&lt;211&gt; 19

&lt;212&gt; PRT

&lt;213&gt; Hepatitis C virus

&lt;400&gt; 60

His Asp Ala Ile Leu His Thr Pro Gly Val Pro Cys Val Arg Glu Gly  
1 5 10 15

Asn Val Ser

&lt;210&gt; 61

&lt;211&gt; 20

&lt;212&gt; PRT

&lt;213&gt; Hepatitis C virus

&lt;400&gt; 61

Cys Val Arg Glu Gly Asn Val Ser Arg Cys Trp Val Ala Met Thr Pro  
1 5 10 15

Thr Val Ala Thr  
20

&lt;210&gt; 62

&lt;211&gt; 20

&lt;212&gt; PRT

&lt;213&gt; Hepatitis C virus

&lt;400&gt; 62

Ala Met Thr Pro Thr Val Ala Thr Arg Asp Gly Lys Leu Pro Ala Thr  
1 5 10 15

Gln Leu Arg Arg  
20

&lt;210&gt; 63

&lt;211&gt; 20

&lt;212&gt; PRT

&lt;213&gt; Hepatitis C virus

&lt;400&gt; 63

Leu Pro Ala Thr Gln Leu Arg Arg His Ile Asp Leu Leu Val Gly Ser  
1 5 10 15

Ala Thr Leu Cys  
20

&lt;210&gt; 64

&lt;211&gt; 20

&lt;212&gt; PRT

&lt;213&gt; Hepatitis C virus

&lt;400&gt; 64

Leu Val Gly Ser Ala Thr Leu Cys Ser Ala Leu Tyr Val Gly Asp Leu  
1 5 10 15

Cys Gly Ser Val  
20

&lt;210&gt; 65

&lt;211&gt; 20

<212> PRT  
<213> Hepatitis C virus

<400> 65

Gln Leu Phe Thr Phe Ser Pro Arg Arg His Trp Thr Thr Gln Gly Cys  
1 5 10 15

Asn Cys Ser Ile

20

<210> 66

<211> 20

<212> PRT

<213> Hepatitis C virus

<400> 66

Thr Gln Gly Cys Asn Cys Ser Ile Tyr Pro Gly His Ile Thr Gly His  
1 5 10 15

Arg Met Ala Trp

20

<210> 67

<211> 20

<212> PRT

<213> Hepatitis C virus

<400> 67

Ile Thr Gly His Arg Met Ala Trp Asp Met Met Met Asn Trp Ser Pro  
1 5 10 15

Thr Ala Ala Leu

20

<210> 68

<211> 20

<212> PRT

<213> Hepatitis C virus

<400> 68

Asn Trp Ser Pro Thr Ala Ala Leu Val Met Ala Gln Leu Leu Arg Ile  
1 5 10 15

Pro Gln Ala Ile

20

<210> 69



<211> 20  
<212> PRT  
<213> Hepatitis C virus

<400> 69

Leu Leu Arg Ile Pro Gln Ala Ile Leu Asp Met Ile Ala Gly Ala His  
1 5 10 15  
Trp Gly Val Leu  
20

<210> 70

<211> 20  
<212> PRT  
<213> Hepatitis C virus

<400> 70

Ala Gly Ala His Trp Gly Val Leu Ala Gly Ile Ala Tyr Phe Ser Met  
1 5 10 15  
Val Gly Asn Met  
20

<210> 71

<211> 20  
<212> PRT  
<213> Hepatitis C virus

<400> 71

Val Val Leu Leu Leu Phe Ala Gly Val Asp Ala Glu Thr Ile Val Ser  
1 5 10 15  
Gly Gly Gln Ala  
20

<210> 72

<211> 20  
<212> PRT  
<213> Hepatitis C virus

<400> 72

Ser Gly Leu Val Ser-Leu Phe-Thr Pro Gly Ala Lys Gln Asn Ile Gln  
1 5 10 15  
Leu Ile Asn Thr

20

&lt;210&gt; 73

&lt;211&gt; 20

&lt;212&gt; PRT

&lt;213&gt; Hepatitis C virus

&lt;400&gt; 73

Gln Asn Ile Gln Leu Ile Asn Thr Asn Gly Gln Trp His Ile Asn Ser  
1 5 10 15

Thr Ala Leu Asn  
20

&lt;210&gt; 74

&lt;211&gt; 21

&lt;212&gt; PRT

&lt;213&gt; Hepatitis C virus

&lt;400&gt; 74

Leu Asn Cys Asn Glu Ser Leu Asn Thr Gly Trp Trp Leu Ala Gly Leu  
1 5 10 15

Ile Tyr Gln His Lys  
20

&lt;210&gt; 75

&lt;211&gt; 20

&lt;212&gt; PRT

&lt;213&gt; Hepatitis C virus

&lt;400&gt; 75

Ala Gly Leu Ile Tyr Gln His Lys Phe Asn Ser Ser Gly Cys Pro Glu  
1 5 10 15

Arg Leu Ala Ser  
20

&lt;210&gt; 76

&lt;211&gt; 20

&lt;212&gt; PRT

&lt;213&gt; Hepatitis C virus

&lt;400&gt; 76

Gly Cys Pro Glu Arg Leu Ala Ser Cys Arg Pro Leu Thr Asp Phe Asp  
1 5 10 15

Gln Gly Trp Gly  
20

&lt;210&gt; 77

&lt;211&gt; 20

&lt;212&gt; PRT

&lt;213&gt; Hepatitis C virus

&lt;400&gt; 77

Thr Asp Phe Asp Gln Gly Trp Gly Pro Ile Ser Tyr Ala Asn Gly Ser  
1 5 10 15

Gly Pro Asp Gln  
20

&lt;210&gt; 78

&lt;211&gt; 20

&lt;212&gt; PRT

&lt;213&gt; Hepatitis C virus

&lt;400&gt; 78

Ala Asn Gly Ser Gly Pro Asp Gln Arg Pro Tyr Cys Trp His Tyr Pro  
1 5 10 15

Pro Lys Pro Cys  
20

&lt;210&gt; 79

&lt;211&gt; 20

&lt;212&gt; PRT

&lt;213&gt; Hepatitis C virus

&lt;400&gt; 79

Trp His Tyr Pro Pro Lys Pro Cys Gly Ile Val Pro Ala Lys Ser Val  
1 5 10 15

Cys Gly Pro Val  
20

&lt;210&gt; 80

&lt;211&gt; 20

&lt;212&gt; PRT

&lt;213&gt; Hepatitis C virus

&lt;400&gt; 80

Ala Lys Ser Val Cys Gly Pro Val Tyr Cys Phe Thr Pro Ser Pro Val  
1 5 10 15

Val Val Gly Thr  
20

<210> 81

<211> 20

<212> PRT

<213> Hepatitis C virus

<400> 81

Pro Ser Pro Val Val Gly Thr Thr Asp Arg Ser Gly Ala Pro Thr  
1 5 10 15

Tyr Ser Trp Gly  
20

<210> 82

<211> 20

<212> PRT

<213> Hepatitis C virus

<400> 82

Gly Ala Pro Thr Tyr Ser Trp Gly Glu Asn Asp Thr Asp Val Phe Val  
1 5 10 15

Leu Asn Asn Thr  
20

<210> 83

<211> 20

<212> PRT

<213> Hepatitis C virus

<400> 83

Gly Asn Trp Phe Gly Cys Thr Trp Met Asn Ser Thr Gly Phe Thr Lys  
1 5 10 15

Val Cys Gly Ala  
20

<210> 84

<211> 20

<212> PRT

<213> Hepatitis C virus

&lt;400&gt; 84

Gly Phe Thr Lys Val Cys Gly Ala Pro Pro Val Cys Ile Gly Gly Ala  
1 5 10 15

Gly Asn Asn Thr  
20

&lt;210&gt; 85

&lt;211&gt; 19

&lt;212&gt; PRT

&lt;213&gt; Hepatitis C virus

&lt;400&gt; 85

Ile Gly Gly Ala Gly Asn Asn Thr Leu His Cys Pro Thr Asp Cys Arg  
1 5 10 15

Lys His Pro

&lt;210&gt; 86

&lt;211&gt; 20

&lt;212&gt; PRT

&lt;213&gt; Hepatitis C virus

&lt;400&gt; 86

Thr Asp Cys Phe Arg Lys His Pro Asp Ala Thr Tyr Ser Arg Cys Gly  
1 5 10 15

Ser Gly Pro Trp  
20

&lt;210&gt; 87

&lt;211&gt; 20

&lt;212&gt; PRT

&lt;213&gt; Hepatitis C virus

&lt;400&gt; 87

Ser Arg Cys Gly Ser Gly Pro Trp Ile Thr Pro Arg Cys Leu Val Asp  
1 5 10 15

Tyr Pro Tyr Arg  
20

&lt;210&gt; 88

&lt;211&gt; 20

&lt;212&gt; PRT

&lt;213&gt; Hepatitis C virus

&lt;400&gt; 88

Cys Leu Val Asp Tyr Pro Tyr Arg Leu Trp His Tyr Pro Cys Thr Ile  
1 5 10 15

Asn Tyr Thr Ile  
20

&lt;210&gt; 89

&lt;211&gt; 20

&lt;212&gt; PRT

&lt;213&gt; Hepatitis C virus

&lt;400&gt; 89

Pro Cys Thr Ile Asn Tyr Thr Ile Phe Lys Ile Arg Met Tyr Val Gly  
1 5 10 15

Gly Val Glu His  
20

&lt;210&gt; 90

&lt;211&gt; 20

&lt;212&gt; PRT

&lt;213&gt; Hepatitis C virus

&lt;400&gt; 90

Met Tyr Val Gly Gly Val Glu His Arg Leu Glu Ala Ala Cys Asn Trp  
1 5 10 15

Thr Pro Gly Glu  
20

&lt;210&gt; 91

&lt;211&gt; 20

&lt;212&gt; PRT

&lt;213&gt; Hepatitis C virus

&lt;400&gt; 91

Ala Cys Asn Trp Thr Pro Gly Glu Arg Cys Asp Leu Glu Asp Arg Asp  
1 5 10 15

Arg Ser Glu Leu  
20

&lt;210&gt; 92

&lt;211&gt; 20

<212> PRT  
<213> Hepatitis C virus

<400> 92

Glu Asp Arg Asp Arg Ser Glu Leu Ser Pro Leu Leu Leu Thr Thr Thr  
1 5 10 15

Gln Trp Gln Val  
20

<210> 93

<211> 9  
<212> PRT  
<213> Hepatitis C virus

<400> 93

Tyr Gln Val Arg Asn Ser Thr Gly Leu  
1 5

<210> 94

<211> 29  
<212> DNA  
<213> Hepatitis C virus

<400> 94

agctaattaa ttaagcttgc atgcctgca

29

<210> 95

<211> 60  
<212> DNA  
<213> Hepatitis C virus

<400> 95

gttaattaac tattagtgat ggtggtgatg gtgtctgccc tcgatcacgt gcaggcctcc

60

<210> 96

<211> 19  
<212> DNA  
<213> Hepatitis C virus

<400> 96

gtttaaccac tgcattgatg

19

<210> 97

<211> 20

<212> DNA  
<213> Hepatitis C virus

<400> 97

gtcccatcga gtgcggctac

20

<210> 98

<211> 45  
<212> DNA  
<213> Hepatitis C virus

<400> 98

cgtgacatgg tacattccgg acacttggcg cacttcataa gcgga

45

<210> 99

<211> 42  
<212> DNA  
<213> Hepatitis C virus

<400> 99

tgcctcatac acaatggagc tctgggacga gtcgttcgtg ac

42

<210> 100

<211> 42  
<212> DNA  
<213> Hepatitis C virus

<400> 100

taccacagcag cgggagctct gttgctcccg aacgcagggc ac

42

<210> 101

<211> 42  
<212> DNA  
<213> Hepatitis C virus

<400> 101

tgtcgtggtg gggacggagg cctgcctagc tgcgagcgtg gg

42

<210> 102

<211> 48  
<212>

<400> 102

cgttatgtgg cccgggtaga ttgagcactg gcagtcctgc accgtctc

48



<210> 103

<211> 42

<212> DNA

<213> Hepatitis C virus

<400> 103

cagggccggtt ctaggcctcc actgcatcat catatcccaa gc

42

<210> 104

<211> 26

<212> DNA

<213> Hepatitis C virus

<400> 104

ccggaatgta ccatgtcacg aacgac

26

<210> 105

<211> 24

<212> DNA

<213> Hepatitis C virus

<400> 105

gctccattgt gtatgaggca gcgg

24

<210> 106

<211> 23

<212> DNA

<213> Hepatitis C virus

<400> 106

gagctcccgc tgctgggtag cgc

23

<210> 107

<211> 25

<212> DNA

<213> Hepatitis C virus

<400> 107

cctccgtccc caccacgaca atacg

25

<210> 108

<211> 27  
<212> DNA  
<213> Hepatitis C virus

<400> 108

ctacccgggc cacataacgg gtcaccg

27

<210> 109

<211> 24  
<212> DNA  
<213> Hepatitis C virus

<400> 109

ggaggcctac aacggccctg gtgg

24

<210> 110

<211> 22  
<212> DNA  
<213> Hepatitis C virus

<400> 110

ttctatcgat taaatagaat tc

22

<210> 111

<211> 23  
<212> DNA  
<213> Hepatitis C virus

<400> 111

gccatacgct cacagccgat ccc

23

<210> 112

<211> 20  
<212> PRT  
<213> Hepatitis C virus

<400> 112

Tyr Glu Val Arg Asn Val Ser Gly Ile Tyr His Val Thr Asn Asp Cys  
1 5 10 15

Ser Asn Ser Ser  
20

<210> 113

<211> 20

<212> PRT

<213> Hepatitis C virus

<400> 113

Thr Asn Asp Cys Ser Asn Ser Ser Ile Val Tyr Glu Ala Ala Asp  
1 5 10 15

Met Ile Met His Thr  
20

<210> 114

<211> 20

<212> PRT

<213> Hepatitis C virus

<400> 114

Ala Ala Asp Met Ile Met His Thr Pro Gly Cys Val Pro Cys Val  
1 5 10 15

Arg Glu Asn Asn Ser  
20

<210> 115

<211> 20

<212> PRT

<213> Hepatitis C virus

<400> 115

Pro Cys Val Arg Glu Asn Asn Ser Ser Arg Cys Trp Val Ala Leu  
1 5 10 15

Thr Pro Thr Leu Ala  
20

<210> 116

<211> 20

<212> PRT

<213> Hepatitis C virus

<400> 116

Val Ala Leu Thr Pro Thr Leu Ala Ala Arg Asn Ala Ser Val Pro  
1 5 10 15

Thr Thr Thr Ile Arg  
20

<210> 117

<211> 20  
<212> PRT  
<213> Hepatitis C virus

<400> 117

Ser Val Pro Thr Thr Thr Ile Arg Arg His Val Asp Leu Leu Val  
1 5 10 15

Gly Ala Ala Ala Phe  
20

<210> 118

<211> 20  
<212> PRT  
<213> Hepatitis C virus

<400> 118

Leu Leu Val Gly Ala Ala Ala Phe Cys Ser Ala Met Tyr Val Gly  
1 5 10 15

Asp Leu Cys Gly Ser  
20

<210> 119

<211> 20  
<212> PRT  
<213> Hepatitis C virus

<400> 119

Tyr Val Gly Asp Leu Cys Gly Ser Val Phe Leu Val Ser Gln Leu  
1 5 10 15

Phe Thr Ile Ser Pro  
20

<210> 120

<211> 20  
<212> PRT  
<213> Hepatitis C virus

<400> 120

Ser Gln Leu Phe Thr Ile Ser Pro Arg Arg His Glu Thr Val Gln  
1 5 10 15

Asp Cys Asn Cys Ser  
20

<210> 121

<211> 20  
<212> PRT

<213> Hepatitis C virus

<400> 121

Thr Val Gln Asp Cys Asn Cys Ser Ile Tyr Pro Gly His Ile Thr  
1 5 10 15

Gly His Arg Met Ala  
20

<210> 122

<211> 20

<212> PRT

<213> Hepatitis C virus

<400> 122

His Ile Thr Gly His Arg Met Ala Trp Asp Met Met Met Asn Trp  
1 5 10 15  
Ser Pro Thr Thr Ala  
20

<210> 123

<211> 135

<212> PRT

<213> hepatitis C virus

<400> 123

Tyr Glu Val Arg Asn Val Ser Gly Met Tyr His Val Thr Asn Asp Cys  
1 5 10 15

Ser Asn Ser Ser Ile Val Tyr Glu Ala Ala Asp Met Ile Met His Thr  
20 25 30

Pro Gly Cys Val Pro Cys Val Arg Glu Asn Asn Ser Ser Arg Cys Trp  
35 40 45

Val Ala Leu Thr Pro Thr Leu Ala Ala Arg Asn Ala Ser Val Pro Thr  
50 55 60

Thr Thr Ile Arg Arg His Val Asp Leu Leu Val Gly Ala Ala Ala Phe  
65 70 75 80

Cys Ser Ala Met Tyr Val Gly Asp Leu Cys Gly Ser Val Phe Leu Val  
85 90 95

Ser Gln Leu Phe Thr Ile Ser Pro Arg Arg His Glu Thr Val Gln Asp  
100 105 110

Cys Asn Cys Ser Ile Tyr Pro Gly His Ile Thr Gly His Arg Met Ala

115

120

125

Trp Asp Met Met Met Asn Trp  
130 135